



HHS Public Access

Author manuscript

Epidemiology. Author manuscript; available in PMC 2019 March 01.

Published in final edited form as:

Epidemiology. 2018 March ; 29(2): 165–174. doi:10.1097/EDE.0000000000000777.

Impact of National Ambient Air Quality Standards nonattainment designations on particulate pollution and health

Corwin M. Zigler*, Christine Choirat*, and Francesca Dominici*

*Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, 02115

Abstract

Background—Despite dramatic air quality improvement in the United States over the past decades, recent years have brought renewed scrutiny and uncertainty surrounding the effectiveness of specific regulatory programs for continuing to improve air quality and public health outcomes.

Methods—We employ causal inference methods and a spatial hierarchical regression model to characterize the extent to which a designation of “nonattainment” with the 1997 National Ambient Air Quality Standard for ambient fine particulate matter (PM_{2.5}) in 2005 causally affected ambient PM_{2.5} and health outcomes among over 10 million Medicare beneficiaries in the Eastern US in 2009–2012.

Results—We found that, on average across all retained study locations, reductions in ambient PM_{2.5} and Medicare health outcomes could not be conclusively attributed to the nonattainment designations against the backdrop of other regional strategies that impacted the entire Eastern United States. A more targeted principal stratification analysis indicates substantial health impacts of the nonattainment designations among the subset of areas where the designations are estimated to have actually reduced ambient PM_{2.5} beyond levels achieved by regional measures, with noteworthy reductions in all-cause mortality, chronic obstructive pulmonary disorder, heart failure, ischemic heart disease, and respiratory tract infections.

Discussion—These findings provide targeted evidence of the effectiveness of local control measures following nonattainment designations for the 1997 PM_{2.5} air quality standard.

INTRODUCTION

While drastic improvements in air quality during the decades following promulgation of the 1970 Clean Air Act (CAA) provided clear indications of the benefits of air quality management [18, 20], recent years’ low levels of ambient pollution have generated renewed uncertainty and controversy surrounding the extent to which control policies continue to improve air quality and health.

Correspondence to: Corwin Zigler, Department of Biostatistics, Building 2, 4th Floor, 655 Huntington Avenue, Boston, MA 02115, 617-432-5014, czigler@hsph.harvard.edu.

Conflicts of Interest: None declared

MS# EDE16-0899

Data sets and code for reproducing the form of the analysis (with simulated Medicare health outcomes data) are available at <https://dataverse.harvard.edu/dataverse/pm97naaqs> and <https://github.com/czigler/PM2.5-Nonattainment>, and are described in supplementary Appendix A.

Existing regulatory health impact assessments such as those carried out by the US Environmental Protection Agency (EPA) typically rely on prospectively modeled air quality predictions under broadly-defined hypothetical scenarios (e.g., assuming full compliance with the 1990 CAA amendments) [6, 21]. To complement these assessments, agencies such as EPA, the National Research Council, and the Health Effects Institute have called for adopting a more performance-oriented approaches that track benefits of specific policies [6, 7, 9, 12, 22, 23]. Direct evaluation of specific strategies using observed data in a manner that is reproducible and less reliant on prospectively-modeled hypothetical air quality scenarios is necessary to strengthen the evidence available to support current policies [2, 6, 25].

The goal of this paper is to use exclusively observed and publicly-available data to characterize the evidence that nonattainment designations for the 1997 National Ambient Air Quality Standard (NAAQS) for particulate matter of $2.5 \mu\text{g}/\text{m}^3$ in diameter or less ($\text{PM}_{2.5}$) caused reductions in ambient $\text{PM}_{2.5}$ and improvements in adverse health outcomes among Medicare beneficiaries living in the Eastern US. Our analysis is rooted in a potential-outcomes perspective on causal inference designed to parse the consequences of the nonattainment designations from changes in air quality and health attributable to other concurrent regulatory actions. This parsing is essential due to the coincident timing of a variety of regional programs that are known to have contributed to improved air quality (see Figure 1). We estimate the effects of localized pollution control measures on ambient $\text{PM}_{2.5}$ and Medicare health outcomes that occurred above and beyond changes due to regional strategies. We provide open-source programs to acquire, link, process, and analyze all data sources. Thus, ours is the first reproducible and open source analysis of the causal effects of NAAQS designations on both ambient air quality and health outcomes.

These efforts towards transparency and reproducibility, in addition to our statistical approach grounded in formal methods for causal inference, are imperative amid the current regulatory climate that subjects existing epidemiological evidence to increased scrutiny amid the prospect of ever-increasing costs of air quality management [3, 18, 25].

METHODS

Nonattainment Designations and Effects of Interest

The NAAQS for $\text{PM}_{2.5}$ were revised in 1997 to introduce annual ($15 \mu\text{g}/\text{m}^3$) and 24-hour ($65 \mu\text{g}/\text{m}^3$) standards for the ambient concentration of $\text{PM}_{2.5}$. After delays due to legal challenges and the need to establish a national monitoring network, counties were officially designated in Spring of 2005 as “nonattainment” of the standard or otherwise “attainment” (meeting or expected to meet the standard) or “unclassifiable” (insufficient data to classify). Such designations were typically based on monitoring data from 2000–2002, States containing counties designated as nonattainment are required to develop or revise State Implementation Plans (SIPs) outlining how a nonattainment area will attain the standards with strategies to reduce ambient concentrations of $\text{PM}_{2.5}$. New control strategies for areas designated as attainment or unclassifiable were not required in SIP revisions. Note that designations occur at the county level, so individual monitoring locations in nonattainment areas may exhibit ambient $\text{PM}_{2.5}$ that does not exceed the NAAQS.

The overall goal is to estimate causal effects of the 2005 nonattainment designations on ambient pollution and Medicare health outcomes in 2010–2012, where causal effects of the designations are defined as the differences between pollution and health outcomes that were observed in the nonattainment areas and those that would have been observed in these same areas had the nonattainment designations never occurred [16]. Importantly, the timing of these designations aligned with various regional or national control strategies (e.g., the Acid Rain Program and Clean Air Interstate Rule) that were expected to reduce ambient PM_{2.5} across the entire Eastern US in both attainment and nonattainment areas. Some nonattainment areas expressly relied on the expected benefits of these regional programs to attain the PM_{2.5} NAAQS. Thus, the present analysis is designed to isolate the incremental benefit of the nonattainment designations above and beyond impacts on pollution and/or health that are attributable to national or regional strategies.

The study design considers years prior to 2005 as the “baseline period,” with a “follow-up” period of 2010–2012 during which the following outcomes are measured: 1) the average annual ambient concentration of PM_{2.5} during 2010–2012; 2) rates of all-cause Medicare mortality (# of deaths per beneficiary) in 2012; and 3) rates of hospitalization (# of hospitalizations per person-year) in 2012 among Medicare beneficiaries for the following conditions: chronic obstructive pulmonary disease (COPD, ICD9-CM 490.xx to 492.xx), cardiovascular stroke (CV stroke, ICD9-CM codes 430.xx-438.xx), heart failure (ICD9-CM codes 428.xx), heart rhythm disorders (ICD9-CM codes 426.xx-427.xx), ischemic heart disease (IHD, ICD9-CM codes 410.xx-414.xx or 429.xx), peripheral vascular disease (ICD9-CM codes 440.xx-448.xx), and respiratory tract infection (ICD9-CM 464.xx to 466.xx and 480.xx to 487.xx). Two hospitalization summaries are constructed and used as covariates in the analysis: one for all respiratory hospitalizations (the sum of COPD and respiratory tract infection admissions) and another for all cardiovascular hospitalizations (All CVD, ICD9-CM codes 390.xx-459.xx, which includes some diagnoses not analyzed individually). Figure 1 describes the timeline of the study period.

Publicly Available and Linked Data Sources

The publicly-available data sources used for the analysis consist of: ambient PM_{2.5} monitoring data from the EPA Air Quality System, nonattainment designations from the EPA Green Book, population demographics from the US Census Bureau (year 2000), county-level smoking rates from previously-used small-area estimation techniques in conjunction with the Center for Diseases Control and Prevention Behavioral Risk Factor Surveillance System [4], climate data from the Automated Surface Observing System, and Medicare health outcomes from the Center for Medicare and Medicaid Services. Medicare mortality data come from the enrollment file, and Medicare Part A claims data were used to identify hospitalizations using principal discharge diagnosis codes. This study was reviewed by the Harvard School of Public Health Institutional Review Board, and details including web links to these data sources appear in eAppendix A.

The above data sets are available at various spatial and temporal resolutions. eAppendix A details how these data were processed and linked to yield the data set used for the present analysis, and also outlines our publicly-available tools to obtain, link, and assemble this data

set. The result is a data set where the units of analysis are 829 pollution monitoring locations, each having measures of ambient PM_{2.5} during the baseline and follow up periods, baseline measures of general population demographics among zip codes with centroids located within 6 miles of the monitoring station, baseline period climate measurements from climate monitors located within 150km of the monitoring site, baseline period smoking rates from the surrounding county, and baseline and follow up period mortality and hospitalization rates from Medicare beneficiaries residing in zip codes with centroids located within 6 miles of the monitoring location. The locations of these monitoring sites appear in Figure 2, 292 of which fall in nonattainment areas and the remaining 537 in areas that were considered in attainment or unclassifiable and will be referred to collectively here as “attainment areas.” Table 1 describes the characteristics of monitoring locations in attainment and nonattainment areas. In sum, the data set contains health-outcomes information on 7,423,025 fee-for-service Medicare enrollees (on whom hospitalization and mortality information is available) and an additional 3,050,558 managed care beneficiaries (on whom only mortality outcomes are available). Note that no areas were added or removed from the set of nonattainment locations until the middle of the follow up period (2011), when some locations switched from nonattainment to attainment with the PM_{2.5} NAAQS, and designations according to the revised 2006 PM_{2.5} standard did not take effect until 2009.

Statistical Methods

The analysis approach is to compare pollution and health outcomes in nonattainment areas against those in attainment areas, properly adjusted for confounding factors that differ between areas. This approach consists of two steps undertaken in sequence.

“Designing” the study with propensity scores

First, *propensity scores* are used to adjust for confounding by grouping attainment and nonattainment locations together on the basis of similarity with regard to baseline characteristics. This is meant to ensure that locations in nonattainment areas are compared only against attainment locations that are similar with regard to the baseline factors in Table 1 [8, 15, 17, 19]. This is accomplished by 1) estimating propensity scores with predicted values from a logistic regression model of the probability that a monitoring location is in a nonattainment area, conditional on the baseline factors noted in Table 1; 2) identifying and omitting locations with features that are not comparable to *any* area in the comparison group (propensity score pruning); and then 3) grouping the remaining locations into four subgroups such that, within each subgroup, attainment and nonattainment areas are similar with regard to all baseline factors in Table 1. Under the crucial assumption that the baseline factors in Table 1 comprise all factors that differ between locations in attainment/nonattainment areas and relate to pollution and health outcomes (i.e., that there are no unmeasured confounders), this propensity score strategy can be regarded as approximating the design of an experiment where nonattainment designations are “randomized” to locations within propensity score groups. Details of the propensity score procedure appear in eAppendix B.

Analysis of causal effects

Following this “design” stage that constructs groups of attainment and nonattainment areas that are comparable with respect to an extensive set of baseline characteristics, causal effects are estimated with the Bayesian approach of [26]. All effects described below are calculated among retained (after pruning) areas actually designated as nonattainment to represent quantities akin to the “Average Treatment Effect on the Treated” [8]. Specifically, a spatial hierarchical regression model for pollution is used to predict the (log-transformed) potential ambient PM_{2.5} concentration in 2010–2012 that would have occurred in nonattainment areas if the designations had never occurred. The model specifies linear adjustment terms for propensity score group indicators and several specific covariates (as detailed in Section), each including an interaction with the indicator of attainment status. The model includes a spatial random effect that acknowledges the similarity of ambient air quality at nearby locations.

For the analysis of each mortality and hospitalization outcome, this same spatial hierarchical pollution model is estimated jointly with a log-linear model for each outcome with the same confounding adjustment for propensity score group and additional covariates. Full details of the model specifications appear in eAppendix C and in [26].

In addition to estimating overall average effects, a secondary objective is to learn the extent to which health impacts of the nonattainment designations differ depending on the extent to which the designations actually improved air quality. This is particularly important in attempt to isolate health effects of the nonattainment designations from those due to concurrent regional strategies. For example, knowledge of differential health effects of nonattainment designations in 1) areas where the designations reduced ambient PM_{2.5} above and beyond reductions due to regional strategies and 2) areas where the designations did not cause additional reductions, would inform whether any health impacts could be specifically attributed to the nonattainment designations. The former group may correspond to locations where designations induced localized control measures, while areas in the latter group may have simply relied on existing regional strategies and not taken any localized action in response to the designation. Towards this goal, we follow the principal stratification approach in [26] to estimate 1) causal effects of the nonattainment designations on health outcomes among areas where the nonattainment designations are estimated to reduce ambient PM_{2.5} by at least 1.0 $\mu\text{g}/\text{m}^3$, called “associative effects,” and 2) causal effects of the nonattainment designations on health outcomes among areas where the nonattainment designations are estimated to not affect ambient PM_{2.5} by more than $\pm 1.0 \mu\text{g}/\text{m}^3$, called “dissociative effects.” In both cases reductions (or lack thereof) represent impacts above and beyond regional strategies. Associative and dissociative effects are estimated with the same models described above. Every attainment and nonattainment location is stratified according to whether the causal effect of the nonattainment designation on ambient PM_{2.5} is estimated as a reduction of at least 1.0 $\mu\text{g}/\text{m}^3$ (for associative effects) or a change smaller than $\pm 1.0 \mu\text{g}/\text{m}^3$ (for dissociative effects), and causal health effects are estimated in each stratum. As with the overall effects on PM_{2.5} and health outcomes, associative and dissociative effect estimates are confined to locations that were actually designated as nonattainment. The strategy for accommodating missing pollution data is described in eAppendix C.3.

RESULTS

The “design” stage that estimates propensity scores and omits (i.e., “prunes”) observations with no comparable location in the opposite attainment group results in a final analysis data set consisting of 404 monitoring locations, each grouped into one of four propensity score subclasses. The population surrounding these retained monitoring locations in 2012 consists of 3,892,984 Medicare fee-for-service beneficiaries and 1,620,778 Medicare managed-care beneficiaries. Figure 2(b) shows the locations of the retained monitors, 183 of which lie in nonattainment areas. Thus, subsequent causal inferences of the effects of the nonattainment designations are confined to the effects on these 183 locations. eTable ?? of the eAppendix provides a descriptive comparison between locations that were pruned and those retained, but note here that locations with the most extreme (high or low) ambient $PM_{2.5}$ in 2002–2004 were among those pruned.

Figure 3 illustrates the success of the propensity score “design” for constructing groups of attainment and nonattainment areas that are comparable with respect to the factors in Table 1. The threat of confounding due to each covariate is quantified with the average standardized mean difference between nonattainment and attainment locations. The red line denoting large differences between attainment and nonattainment locations in the full sample of 829 locations indicates a strong threat of confounding before applying any propensity score adjustment. In contrast, the blue line displays average differences within propensity score subclass, which are close to zero for all covariates indicating adequate balance and a reduced threat of confounding. While overall balance is achieved on average across the four subgroups, balance is not perfect for all covariates, and those exhibiting imbalance even after the propensity score subgrouping (denoted in Table 1), as well as baseline $PM_{2.5}$ in 2002–2004 are included in the statistical models as covariates for adjustment within propensity score group [19].

Overall Average Causal Effect Estimates

The estimated average causal effect of the nonattainment designations on average ambient $PM_{2.5}$ in 2010–2012 among the 183 nonattainment areas in the pruned sample is $-0.339 \mu g/m^3$, with 95% posterior interval $(-1.542, 0.688)$, providing limited evidence that, on average across all 183 locations, the nonattainment designations causally improved this measure of air quality above and beyond improvements due to regional strategies impacting both attainment and nonattainment areas.

Figure 4 summarizes posterior distributions of the average causal effects of the nonattainment designations on the Medicare mortality rate (deaths per 1000 beneficiaries), and hospitalization rates (hospitalizations per 1000 person-years) among nonattainment areas. For all-cause mortality, the posterior mean estimates indicate that, on average across nonattainment areas, the nonattainment designations reduced the mortality rate by 1.251, with 95% posterior interval $(-2.631, 0.108)$. Point estimates for the overall average effects of the designations on hospitalization outcomes ranged from an average reduction of 0.651 for CV stroke, to an increase of 0.440 for respiratory tract infections, although 95% posterior intervals for all hospitalization outcomes included 0 and were unable to rule out the possibility of no average effect.

Average Associative and Dissociative Effects

Estimates of average associative and dissociative effects are depicted alongside overall average health effects in Figure 4. Across all health outcomes, dissociative effects are estimated to be very similar to overall health effects, with posterior intervals overlapping 0. This implies that, among areas where the nonattainment designations did not substantially impact $PM_{2.5}$ over what was accomplished by regional strategies, there is no evidence that the nonattainment designations improved health outcomes. The average number of nonattainment locations estimated to have exhibited no substantial impact on $PM_{2.5}$ (i.e., $< 1.0 \mu g/m^3$) varied from 50.7 to 61.4, depending on the outcome.

In contrast, average associative effects are estimated to be different from overall average and average dissociative effects for many outcomes. In particular, the average associative effects for all-cause mortality, COPD, Heart Failure, IHD, and respiratory tract infections rates are estimated to be -3.161 ($-5.188, -1.210$), -2.658 ($-4.305, -1.167$), -2.375 ($-4.349, -0.436$), -2.597 ($-4.242, -1.137$), and -2.611 ($-4.440, -0.840$), respectively. This provides evidence that the nonattainment designations causally reduced rates of these health outcomes among areas where the designations causally reduced ambient $PM_{2.5}$. All other associative effects are estimated with 95% posterior intervals including zero, with point estimates suggesting increases in rates of CV stroke and heard rhythm disorders. The average number of nonattainment locations estimated to have exhibited causal reductions in ambient $PM_{2.5}$ (of at least $1.0 \mu g/m^3$) varied from 59.7 to 75.3, depending on the outcome.

A sensitivity analysis analogous to that above but without the propensity score pruning (i.e., using all 829 locations) exhibits poorer covariate balance and more pronounced overall estimates of pollution and health improvements, including significant reductions in all-cause mortality and heart failure. Results for associative and dissociative effects are similar to those presented here. Details appear in eAppendix D.

DISCUSSION

The multifaceted system of air quality management in the US has generated enormous benefits at substantial cost, with regulations targeting $PM_{2.5}$ alone constituting over half of the benefits and nearly half of the costs of all monetized federal regulations [13]. Periodic evaluation of the effectiveness of specific components of this system is central to its refinement. Towards this goal, we have provided a focused, empirical, and reproducible evaluation of enforcing NAAQS for $PM_{2.5}$ with nonattainment designations.

We employed a causal inference perspective designed to directly answer whether ambient $PM_{2.5}$ and Medicare health outcomes among nonattainment locations were improved in 2010–2012, relative to what would have occurred if EPA had never prompted areas to take local control actions with the designations. Our approach uses observed pollution and health outcomes, making use of modern tools for causal inference and confounding adjustment to learn about the counterfactual scenario where nonattainment designations were never instituted. We used principal stratification to determine whether causal health effects of the designations differed among areas where the designations did and did not actually improve air quality above and beyond improvements due to regional strategies.

Importantly, the analysis does not estimate the entire benefit of setting and enforcing the 1997 NAAQS for PM_{2.5}. Several national or regional control programs were designed in part to help achieve the NAAQS but actually affected attainment and nonattainment areas alike. In particular, programs to reduce important PM_{2.5} precursors such as the Acid Rain Program and the Clean Air Interstate Rule, as well as regulations on diesel emissions, were widely recognized as important for achievement of the NAAQS. The impact of these programs across both attainment and nonattainment areas is evident from Table 1; average ambient PM_{2.5} went down from 11.59 during the baseline period to 9.39 during the follow up period among attainment areas, and the analogous decrease in nonattainment areas was from 14.48 to 11.13. Note that random regression to the mean for areas with particularly high (or low) PM_{2.5} is possible, but not expected to play a large role with a measure as temporally stable as multi-year annual averages of PM_{2.5}.

The present analysis regarding monitoring locations in nonattainment areas as “treated” by the designation is designed to estimate the extent to which any localized pollution control measures implemented in nonattainment areas improved pollution and health outcomes above and beyond those conferred by regional programs. Importantly, not all nonattainment areas were subject to controls on local sources, in particular when EPA recognized that the expected benefits of regional controls would bring an area into attainment without further local action. What’s more, monitoring locations that are in nonattainment areas but don’t themselves indicate elevated PM_{2.5} are likely not the target of localized actions [1]. Thus, the diversity of local action (or inaction) in response to the nonattainment designations renders the analysis akin to an “intent-to-treat” analysis in a clinical study that estimates the effect of *assignment* to an intervention, as opposed to actual receipt of that intervention. This diversity of local action, in combination with the recognized importance of regional programs across the entire Eastern US, likely explains the result that overall average effects on ambient PM_{2.5} and on health outcomes were not significantly different from zero. This result is consistent with previous analyses of nonattainment designations that similarly average over locations that may or may not have been subject to actual pollution control measures [5].

The diversity of actions that may or may not follow a nonattainment designation also underscores the importance of the principal stratification analysis that compares health effect estimates among locations where ambient PM_{2.5} was actually reduced by the nonattainment designation (associative effects) against those where pollution was not measurably affected (dissociative effects); a distinction that presumably corresponds to whether local PM_{2.5} control measures were actually implemented in the area. This analysis indicated that the designations did not causally impact health outcomes among locations where the designations did not meaningfully affect ambient PM_{2.5} in 2010–2012, that is, average dissociative effects were all estimated near zero. In contrast, average associative effects were estimated to suggest causal reductions in rates of all-cause Medicare mortality and hospitalization for COPD, heart failure, IHD, and respiratory tract infections among locations where the designations decreased ambient PM_{2.5} by at least 1.0 $\mu\text{g}/\text{m}^3$. Taken together, these results imply that effects on these health outcomes occurred in conjunction with improvements to ambient PM_{2.5} attributable to the designations themselves (i.e., not due to regional programs). These results are consistent with the simple conclusion that

PM_{2.5} reductions cause health improvements and, by extension, with the presumed causal pathway whereby reducing ambient concentrations of PM_{2.5} in turn causes health improvements. However, the principal stratification analysis cannot definitively support the conclusion that health improvements are due to reductions in average ambient PM_{2.5} in 2010–2012, as would be investigated in a formal “mediation analysis” [11, 14]. Principal stratification in this context cannot rule out, for example, the possibility that the concurrence between nonattainment-induced reductions in PM_{2.5} and the nonattainment-induced improvements in health outcomes is actually due to some other causal pathway (e.g., a different measure of PM_{2.5} over a different time frame) or some other factor common to the areas where both PM_{2.5} and health were improved (e.g., economic changes associated with causal improvements in both pollution and health) [10, 24].

The prospect of confounding presents a key challenge to the analysis of such epidemiological data, particularly so in this instance because nonattainment areas were designated precisely because they exhibited poor air quality. Poor air quality may be associated with a multitude of factors that differentiate attainment and nonattainment areas and may also relate to health outcomes. The propensity score strategy was able to group together attainment and nonattainment locations that were similar on the basis of baseline pollution levels, characteristics of the population residing near the monitoring locations, weather factors, and smoking patterns, thus minimizing the chance of confounding with regard to these factors. Importantly, the primary analysis discarded monitor locations in nonattainment areas that were not similar to *any* attainment area on the basis of observed confounders (and vice versa), restricting conclusions to only the retained subset of 183 monitoring locations located in nonattainment areas. Sensitivity analyses indicated that results without this pruning indicated even more pronounced overall effects of the nonattainment designations, but these should be interpreted in light of the extrapolation across areas that may not be comparable. While the analysis strategy adopted here is specifically designed to mitigate bias due to measured confounders, the prospect of unmeasured confounding remains a threat to the validity of results. If there exist unmeasured factors related to pollution and/or health outcomes that, even after adjustment for all observed factors in Table 1, still differ between attainment and nonattainment areas, then the results are subject to unmeasured confounding.

Nonattainment designations are one key mechanism for air quality management in the US, and represent a key step to the implementation of the NAAQS. Evaluating the effectiveness of these designations required construction of a complex data resource that links together information on regulatory actions, ambient air quality, population characteristics, and health information on the entire US Medicare population in the Eastern US. A key feature of this research is the provision of open-source tools to obtain, link, and analyze these publicly-available data sets (see eAppendix A for details) to promote transparency and reproducibility of results, as well as to foster similar investigations of other regulatory interventions. Despite the limitations of the present analysis, it provides evidence of the effectiveness of one integral feature of air quality management in the US, and represents a distinct perspective that should be interpreted in conjunction with - not instead of - the large body of epidemiological research motivating the setting and implementation of NAAQS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support: This work was supported by research funding from NIHR01ES026217, R01GM111339, R01ES024332, and P50MD010428, EPA 83587201, and HEI 4909 and 4953. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the USEPA. Further, USEPA does not endorse the purchase of any commercial products or services mentioned in the publication.

The authors acknowledge and thank John Bachmann, Lauren Hund, and Yun Wang for their contributions to this work.

References

1. Auffhammer, Maximilian, Bento, Antonio M., Lowe, Scott E. Measuring the effects of the Clean Air Act Amendments on ambient PM10 concentrations: The critical importance of a spatially disaggregated analysis. *Journal of Environmental Economics and Management*. 2009; 58:15–26. author (collaboration).
2. Benneer, Lori Snyder, Coglianese, Cary. Measuring Progress: Program Evaluation of Environmental Policies. *Environment: Science and Policy for Sustainable Development*. 2005; 47:22–39.
3. Dominici, Francesca, Greenstone, Michael, Sunstein, Cass R. Particulate Matter Matters. *Science*. 2014; 344:257–259. [PubMed: 24744361]
4. Dwyer-Lindgren, Laura, Mokdad, Ali H., Srebotnjak, Tanja, Flaxman, Abraham D., Hansen, Gillian M., Murray, Christopher JL. Cigarette smoking prevalence in US counties: 1996–2012. *Population Health Metrics*. 2014; 12:5. [PubMed: 24661401]
5. Greenstone, Michael. Did the Clean Air Act cause the remarkable decline in sulfur dioxide concentrations? *Journal of Environmental Economics and Management*. 2004; 47:585–611.
6. HEI Accountability Working Group. Assessing the Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research. Communication 11. Boston, MA: Health Effects Institute; 2003.
7. Hidy, GM., Brook, JR., Demerjian, KL., Pennell, WT., Scheffe, RD. Technical Challenges of Multipollutant Air Quality Management. New York, NY: Springer; 2011.
8. Ho, Daniel E., Imai, Kosuke, King, Gary, Stuart, Elizabeth A. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*. 2007; 15:199–236.
9. Hubbell, Bryan. Assessing the Results of Air Quality Management Programs. *EM Magazine*. 2012:8–15.
10. Joffe, Marshall M., Greene, Tom. Related Causal Frameworks for Surrogate Outcomes. *Biometrics*. 2009; 65:530–538. [PubMed: 18759836]
11. MacKinnon, David Peter. Introduction to statistical mediation analysis. Routledge; 2008.
12. National Research Council. Air Quality Management in the United States. Washington, DC: National Academies Press; 2004.
13. Office of Management and Budget. Tech rep. OMB; Washington, DC: 2013. 2013 Draft report to Congress on the benefits and costs of federal regulation and unfunded mandates on state, local, and tribal entities.
14. Robins, James M., Greenland, Sander. Identifiability and Exchangeability for Direct and Indirect Effects. *Epidemiology*. 1992; 3:143–155. [PubMed: 1576220]
15. Rosenbaum, Paul R., Rubin, Donald B. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. 1983; 70:41–55.
16. Rubin, Donald B. Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*. 1978; 6:34–58.

17. Rubin, Donald B. For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*. 2008; 2:808–840.
18. Samet, Jonathan M. The Clean Air Act and Health A Clearer View from 2011. *New England Journal of Medicine*. 2011; 365:198–201. [PubMed: 21732828]
19. Stuart, Elizabeth A. Matching methods for causal inference: A review and a look forward. *Statistical science : a review journal of the Institute of Mathematical Statistics*. 2010; 25:1–21. [PubMed: 20871802]
20. U.S. EPA. Integrated Science Assessment for Particulate Matter. Research Triangle Park, NC: U.S. Environmental Protection Agency Office of Research and Development; 2009.
21. U.S. EPA. Revised Draft Report. US EPA Office of Air and Radiation; Research Triangle Park, NC: 2010. The Benefits and Costs of the Clean Air Act: 1990 to 2020.
22. U.S. EPA. Designing Research to Assess Air Quality and Health Outcomes from Air Pollution Regulations. Research Triangle Park, NC: 2013. Workshop on Designing Research to Assess Air Quality and Health Outcomes from Air Pollution Regulations.
23. van Erp, Annemoon M., Kelly, Frank J., Demerjian, Kenneth L., Pope, C Arden, III, Cohen, Aaron J. Progress in research to assess the effectiveness of air quality interventions towards improving public health. *Air Quality, Atmosphere & Health*. 2012; 5:217–230.
24. Vander Weele, Tyler J. Marginal Structural Models for the Estimation of Direct and Indirect Effects. *Epidemiology*. 2009; 20:18–26. [PubMed: 19234398]
25. Zigler, Corwin Matthew, Dominici, Francesca. Point: Clarifying Policy Evidence With Potential- Outcomes Thinking Beyond Exposure-Response Estimation in Air Pollution Epidemiology. *American Journal of Epidemiology*. 2014; 180:1133–1140. [PubMed: 25399414]
26. Zigler, Corwin M., Dominici, Francesca, Wang, Yun. Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes. *Biostatistics (Oxford, England)*. 2012; 13:289–302.

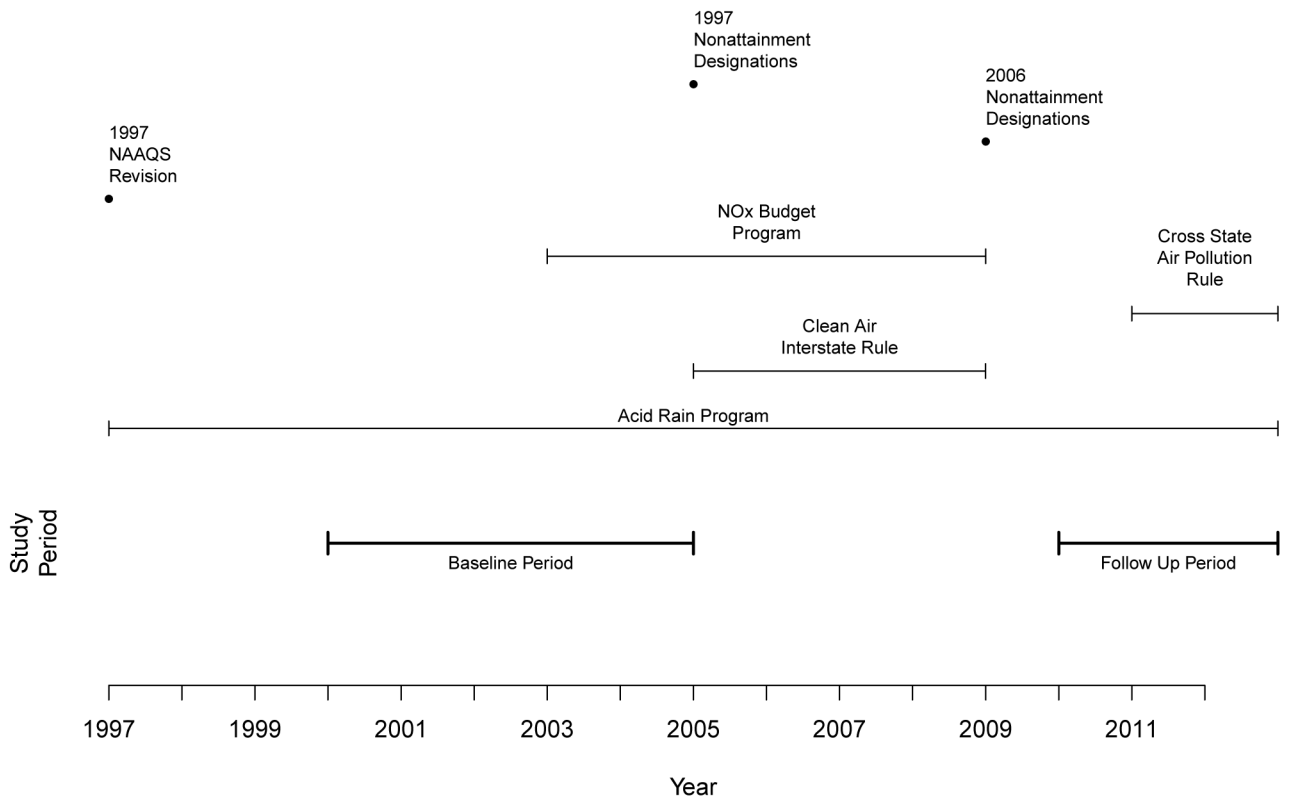


Figure 1. Study timeline depicting 1997 National Ambient Air Quality Standards revisions, nonattainment designations (in 2005), and other concurrent regulatory programs.

(a) All Available Monitors



(b) Retained Monitors After Propensity Score Pruning



Figure 2. Locations of all (a) the 829 PM_{2.5} monitoring locations available for the initial analysis set and (b) the 404 locations retained after propensity score pruning.

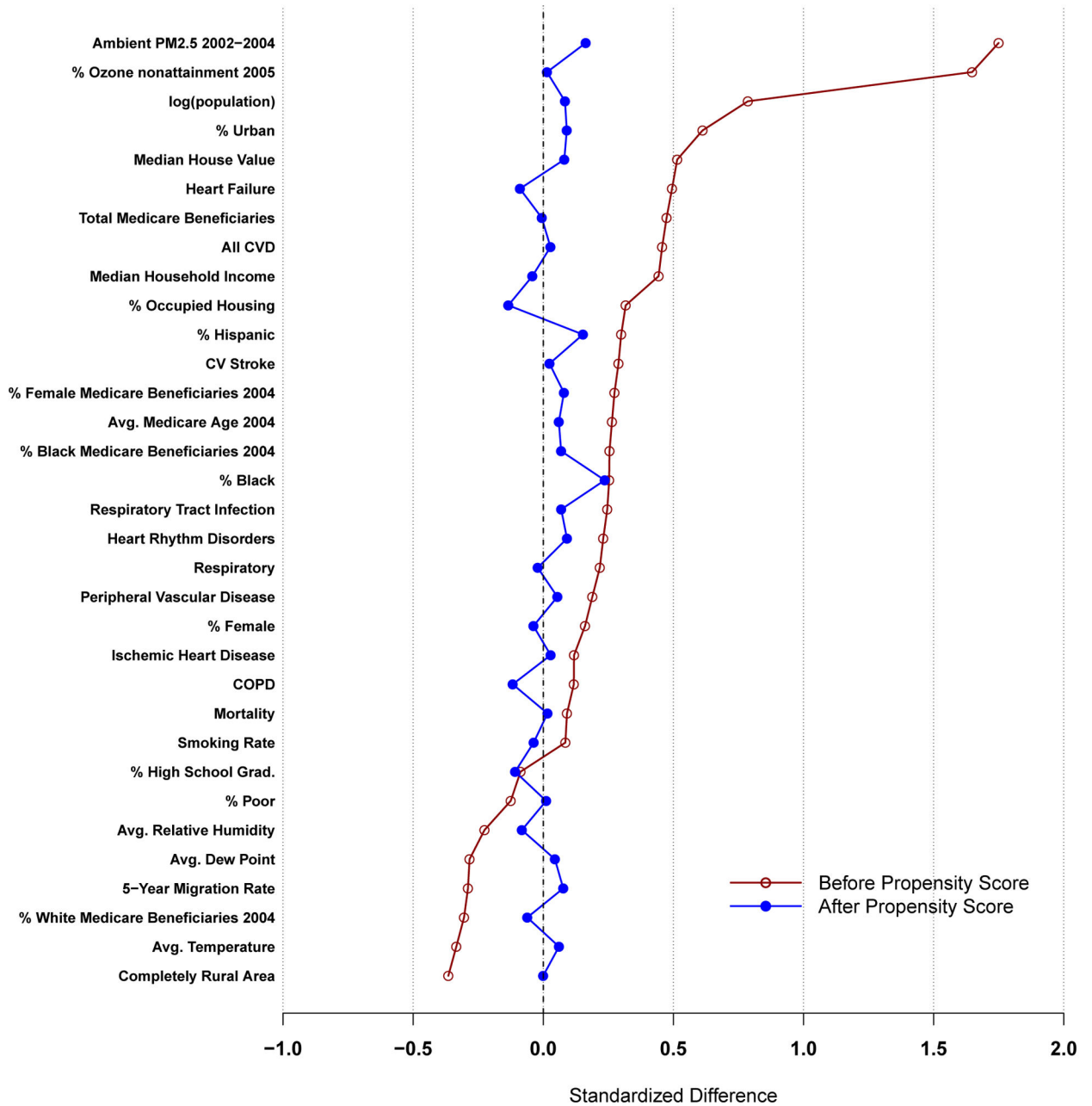


Figure 3.

Plots of average standardized mean difference in each measured covariate between attainment and nonattainment locations. Red line denotes differences before propensity score pruning and subclassification, blue line denotes differences within propensity score subclass. Values closer to zero represent covariate balance between attainment and nonattainment areas. CV(D) is cardiovascular (disease). COPD is chronic obstructive pulmonary disease.

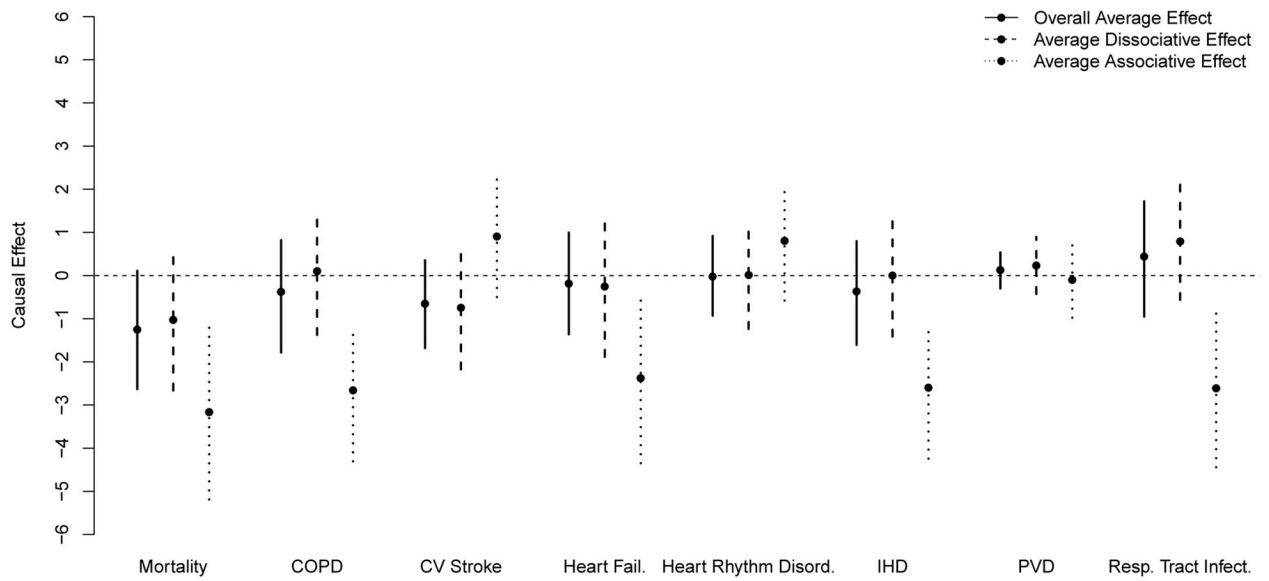


Figure 4. Posterior mean point estimates and 95% posterior probability intervals for overall, associative, and dissociative effects in the analysis of PM_{2.5} nonattainment designations. Effects are measured as risk differences. Overall average effects are calculated across all 183 nonattainment locations. Dissociative effects are averaged across locations where PM_{2.5} is estimated to have changed less than 1 $\mu\text{g}/\text{m}^3$. Associative effects are averaged across locations where PM_{2.5} is estimated to have been reduced by at least 1 $\mu\text{g}/\text{m}^3$.

Table 1
Summary statistics for covariates and outcomes available for the analysis of PM_{2.5} nonattainment designations.

	Attainment Areas (n = 537)		Nonattainment Areas (n = 292)	
	Mean	SD	Mean	SD
<u>Pollution Monitoring Data</u>				
Ambient PM _{2.5} 2002–2004 ($\mu\text{g}/\text{m}^3$) ^{ab}	11.59	1.88	14.48	1.39
% Ozone nonattainment 2005 ^a	30.54	46.1	92.47	26.44
<u>Population Demographics (Year 2000)</u>				
log(population) ^{ab}	11.42	1.59	12.59	1.38
Completely Rural Area ^a	0.08	0.27	0.01	0.08
% Urban ^{ab}	76.64	29.15	91.15	16.55
% Black ^b	17.01	17.58	21.7	19.39
% Hispanic ^b	4.37	5.86	6.5	8.18
% HS Grad. ^{ab}	31.17	6.74	30.56	7.27
Median HH Inc. (\$) ^a	37336.79	8862.03	42596.43	14248.56
% Poor ^{ab}	14.43	5.89	13.65	6.57
% Female	51.49	1.47	51.73	1.49
% Occupied Housing ^{ab}	90.65	7.2	92.42	3.19
5-Year Migration Rate ^a	0.49	0.08	0.47	0.07
Median House Value (\$) ^a	98235.36	38762.8	129863.19	77754.06
Smoking Rate ^a	0.26	0.03	0.26	0.03
<u>Climate (Years 2004 – 2006)</u>				
Avg. Dew Point (°F) ^a	46.34	7.95	44.63	3.15
Avg. Temperature (°F) ^a	56.14	8.53	53.96	3.57
Avg. Rel. Humidity (%) ^{ab}	72.31	1.84	71.92	1.62
<u>Baseline Medicare Characteristics (Year 2004)</u>				
Total Medicare Benef. ^{ab}	9305.46	10133.62	15867.97	16752.38
Avg. Medicare Age (years) ^{ab}	75.19	0.94	75.45	1.01
% Female Medicare Benef. ^{ab}	59.09	2.91	59.81	2.28
% White Medicare Benef. ^{ab}	87.13	15.43	81.45	21.44

	Attainment Areas (<i>n</i> = 537)		Nonattainment Areas (<i>n</i> = 292)	
	Mean	SD	Mean	SD
% Black Medicare Benef. ^{a,b}	10.76	14.6	15.26	20.32
Mortality ^a	52.6	9.17	53.37	7.54
All CVD ^a	108.69	28.45	120.9	27.55
Respiratory ^a	31.4	13.68	34.36	11.96
COPD	10.4	6.96	11.16	5.83
CV Stroke	18.53	6.93	20.65	5.93
Heart Failure	23.66	8.98	28.51	10.57
Heart Rhythm Disorders	15.08	5.46	16.07	4.58
Ischemic Heart Disease	29.8	11.31	30.75	9.08
Peripheral Vascular Disease	6.86	4.54	7.18	2.74
Respiratory Tract Infection	21	8.7	23.2	7.87
<u>Pollution and Health Outcomes</u>				
Ambient PM _{2.5} 2010–2012 ($\mu\text{g}/\text{m}^3$)	9.39	1.65	11.13	1.35
Mortality 2012	48.04	7.5	48.35	9.51
COPD 2012	10.07	6.7	11.42	5.47
CV Stroke 2012	14.71	6.32	15.93	4.97
Heart Failure 2012	16.32	7.82	18.99	7.42
Heart Rhythm Disorders 2012	13.65	5.31	14.86	4.49
Ischemic Heart Disease 2012	15.64	6.82	15.72	5.95
Peripheral Vascular Disease 2012	4.19	2.33	4.7	2.57
Respiratory Tract Infection 2012	15.15	7.52	16.26	6.08

Variables marked with ^a are those included in the model that estimates the propensity score, and those marked with ^b are those included for additional covariate adjustment in models for pollution and health outcomes. Medicare health outcomes are listed as rates per 1000 beneficiaries for mortality and per 1000 person-years for hospitalizations.