The Threat to Air Pollution Health Studies Behind the Environmental Protection Agency's Cloak of Science Transparency

See also the *AJPH* Environmental Health Workforce & Regulation section, pp. 284–298.

This year marks the 50th anniversary of the first Earth Day (April 22), the Clean Air Act of 1970, and the creation of the Environmental Protection Agency (EPA). Over those five decades, we have seen remarkable improvements in air quality across the United States and significant improvements in public health. Indeed the Office of Management and Budget reports that regulation of fine particle ($PM_{2.5}$) air pollution is the most cost effective of all federal regulations.[1] This success in controlling air pollution should be celebrated as a public health triumph. Thus, the EPA's proposal to restrict the use of key scientific evidence regarding the health effects of air pollution under the guise of "scientific transparency" is disturbing.[2]

The EPA's "Strengthening Transparency in Regulatory Science" rule is the ratification of efforts by special interests over more than two decades to undermine the science underpinning air pollution regulations. Twenty-five years ago, we published with our colleagues results of two prospective cohort studies, the Harvard Six Cities Study[3] and the American Cancer Society Study,[4] reporting that mortality risk increased linearly with long-term exposure to $PM_{2.5}$ air pollution. Observed $PM_{2.5}$–mortality associations were remarkably robust, especially for cardiopulmonary mortality. Furthermore, associations were much larger than expected based on previous short-term associations observed in daily time series studies.

Under common scientific practice, publication of these results would lead to attempts to verify and replicate these results using studies in independent populations by independent investigators. However, because the EPA cited these studies as key in setting air quality standards for $PM_{2.5}$, there were calls for examination of the validity, analytic methods, and data quality of the original studies and demands for access to the raw data.

We support principles of data accessibility and transparency in conducting and reporting scientific research. We also respect and adhere to ethical and legal obligations to protect the private and confidential data of study participants, including guarantees of confidentiality given to each participant, agencies providing mortality data, and institutional review boards (human studies committees) overseeing these studies. Nevertheless, given the public health and policy importance of these studies, investigators of both studies agreed to provide cohort and related data for an unrestricted, intensive reanalysis by an independent research team selected and overseen by the Health Effects Institute with full assurance of confidentiality of participant information. These independent reviewers conducted data quality audits, evaluated the reproducibility of the originally published findings, and assessed the sensitivity of the analyses to alternative methods and additional data. After a three-year effort, they published a 300-page report that found the data were of good quality, the original analyses could be reproduced, and the results were relatively unaffected by alternative analyses.[5]

Although this reanalysis was reassuring, it does not provide the scientific validation that comes from replicating these results by independent investigators in independent cohorts. In the 25 years since these two seminal studies were published, there have been dozens of additional longitudinal cohort studies published using independent cohorts from the United States, Canada, Europe, and Asia, confirming and building on the $PM_{2.5}$–mortality association.[6]

Despite these efforts, the American Cancer Society and Harvard Six Cities studies continue to be characterized by special interest groups as "secret science." This characterization is simply false on its face. The methodology, protocol, and results of these studies have been published in high-quality peer-reviewed scientific journals. We provided full data access for audits and independent reanalyses. We have continued to be actively involved in open, collaborative, extended analysis efforts in a way that contributes to scientific understanding and that does not violate commitments to the privacy and confidentiality of research participants.

Prominent scientific organizations, editors of major scientific journals, and others have affirmed that some of the most important and informative studies in medicine generally, and environmental medicine particularly, include private and confidential data that cannot be fully made public.[7] They have noted the imprudence of discounting or disallowing the use of results from these studies in science-based public policy. Special interest groups still demand the release of the individual data and argue that without such release these studies should not be used to inform public policy.

Why can't we just anonymize the data, as is standard practice in clinical trials? In a clinical trial, essential individual data include age, sex, race, and an indicator of randomized treatment group. Participants can feel safe that they and their records cannot be identified. In environmental epidemiology, the essential

ABOUT THE AUTHORS

Douglas W. Dockery is with the Departments of Environmental Health and Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA. C. Arden Pope III is with the Department of Economics, Brigham Young University, Provo, UT.

Correspondence should be sent to Douglas W. Dockery, John L. Loeb and Frances Lehman Loeb Research Professor of Environmental Epidemiology, Harvard T.H. Chan School of Public Health, 665 Huntington Ave, Boston, MA 02115 (e-mail: ddockery@hsph.harvard.edu). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.

This editorial was accepted December 14, 2019.

doi: 10.2105/AJPH.2019.305531

individual data likewise include age, sex, race, and environmental exposure. However, environmental exposure is not randomly assigned. Rather a study participant's environment is defined by where they live, that is the air, the water, and the land at or near their residential location.

Consider $PM_{2.5}$ air pollution that can now be modeled and mapped across the United States by metropolitan areas, counties, zip codes, census tracts, one-kilometer square grids, and even geocoded residential addresses. Would people be willing to participate in a public health study if they understood that they could likely be identified based on their age, sex, race, and where they live? Even the US National Center for Health Statistics, which provides de-identified mortality and national health interview survey data, severely restricts the use and reporting of data linked to geographic information, because of the high risks of compromising required and assured confidentiality.

Further consider how environmental epidemiology informs public health policy. In evaluating therapeutic interventions, clinical trials provide evaluation of safety and efficacy before release to the public. In addition, retrospective postmarketing surveillance is required to detect unexpected adverse effects. For environmental contaminants, there is no prerelease evaluation of effects in populations who are exposed. Primary evidence of adverse effects comes from environmental epidemiology studies, often looking retrospectively at previous exposures based on residential location. That is, environmental epidemiology provides a critical scientific approach for detecting adverse effects of contaminants in the environment.

The EPA transparency rule would not only preclude consideration of previous studies but also hamstring future epidemiologic studies of hazards based on residential location, including air pollution, water contamination, hazardous waste, radiation, spills and accidental releases, and other known or unidentified hazards. The effect not only suppresses previous inconvenient evidence but also discourages participation in future epidemiologic studies. If potential participants are discouraged from participating in population studies, opponents of environmental regulation will have succeeded in removing a key scientific approach to detect and quantify hazards.

The EPA rule, although appearing to formalize good scientific practice under the cloak of scientific transparency, would severely limit the use of epidemiology for surveillance of the deleterious health effects of contaminants in the environment and surreptitiously discourage participation in studies that address environmental issues.

Imagine the effect of a rule that required release of approximate residential addresses of participants in clinical trials or in other studies that use medical records and other clinical-based data. The impact would be devastating to the conduct of these studies and in advancing knowledge and practice in clinic medicine. Imagine further a rule that similarly required release of approximate addresses of participants in postmarketing surveillance epidemiologic studies. The ability to detect and monitor unexpected adverse effects and interactions would be severely compromised. Why does the EPA propose to restrict our scientific ability to detect the adverse effects of environmental contaminants under the cloak of scientific transparency? AJPH

Douglas W. Dockery, ScD
C. Arden Pope III, PhD

CONTRIBUTORS

The authors contributed equally to this article.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Office of Information and Regulatory Affairs. *2015 Report to Congress on the Benefits and Costs of Federal Regulations and Agency Compliance With the Unfunded Mandates Reform Act*. US Office of Management and Budget. 2016; Available at https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/inforeg/inforeg/2015_cb/2015-cost-benefit-report.pdf. Accessed December 22, 2019.

2. US Environmental Protection Agency. Strengthening transparency in regulatory science: a proposed rule by the Environmental Protection Agency on 04/30/2018. *Fed Regist*. 2018;83(83):18768–18774. 40 CFR Part 30.

3. Dockery DW, Pope CA 3rd, Xu X, et al. An association between air pollution and mortality in six US cities. *N Engl J Med*. 1993;329(24):1753–1759.

4. Pope CA 3rd, Thun MJ, Namboodiri MM, et al. Particulate air pollution as a predictor of mortality in a prospective study of US adults. *Am J Respir Crit Care Med*. 1995;151(3 pt 1):669–674.

5. Krewski D, Burnett RT, Goldberg MS, et al. *Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality*. Cambridge, MA: Health Effects Institute; 2000.

6. Pope CA, Coleman N, Pond ZA, Burnett RT. Fine particulate air pollution and human mortality: 25+ years of cohort studies. *Environ Res*. 2019; Epub ahead of print.

7. Thorp HH, Skipper M, Kiermer V, Berenbaum M, Sweet D, Horton R. Joint statement on EPA proposed rule and public availability of data (2019). *Science* 2019;366(6470):eaba3197.