POLICY FORUM

SCIENCE AND REGULATION

Don't abandon evidence and process on air pollution policy

Who decides how to establish causality?

By **Gretchen T. Goldman**[1] *and* **Francesca Dominici**[2]

Air pollution kills—scientists have known this for many years. But how do they know? The global scientific community has developed and agreed upon a framework that draws on multiple lines of evidence across different scientific disciplines to assess the existence and strength of links between air pollution and health. In the United States, federal policies require use of this science-based framework to ensure that air pollution standards protect the public's health. But now this science-based policy process—and public health—are at risk. Recent developments at the U.S. Environmental Protection Agency (EPA) stand to quietly upend the time-tested and scientifically backed process the agency relies on to protect the public from ambient air pollution (*1*). One of these developments—changes in how the EPA handles causality between air pollutants and health effects—has received less attention but, if enacted, would alter the approach that the EPA has used for more than a decade to set health-based air pollutant standards. At the March meeting of the EPA's Clean Air Scientific Advisory Committee (CASAC) (*2*), these changes may begin to unfold. The agency now faces a dilemma. If the EPA leadership embraces the process proposed by the current CASAC chair, it will fundamentally change the EPA's process for scientific assessment. If the EPA leadership ignores the CASAC recommendations, then the agency would be declining to listen to (what should be) its top science advisers, thus eroding the foundational concept of peer review as central to ensuring the use of strong science in policy decisions

WEIGHT OF THE EVIDENCE

Consistent with how the broader scientific community builds consensus on a topic, the EPA for decades has methodically assessed the strength of the relationship between air pollution and health outcomes, and has determined the need for strengthening pollutant protections. These determinations have been made only after robust, transparent peer review with public input. The Clean Air Act–mandated CASAC, a group of experts that operates independently from the EPA, has provided science advice on ambient air pollutant standards since the law's enactment. Their input is supplemented by pollutant-specific panels of experts that span scientific disciplines and have long histories of peer-reviewed publications. These review panels provide pollutant-specific, evidence-based advice needed for EPA to set the air pollution standards. Even in the face of enormous political and financial pressures to roll back pollution controls, this process has worked remarkably well across both Republican and Democratic administrations and has been upheld in the courts, where several legal challenges to its use in past pollutant reviews have been defeated [see supplementary materials (SM), section 1]. Political decisions haven't always aligned with the science, but the process for developing and communicating policy-relevant scientific assessments has remained largely intact (*3*).

Within these scientific assessments, the EPA has applied a weight-of-the-evidence approach for causality determination using a five-level hierarchy, ranging from a "causal relationship" to "no evidence of a causal relationship," to assess links between air pollutants and health effects. This approach is rooted in the scientific community's decades-long effort to evaluate the relationship between cause and effect, beginning with work by Sir Bradford Hill in 1965 and a 1964 report from the U.S. Surgeon General, and then with approaches later developed by leading scientific bodies such as the National Academy of Medicine and International Agency for Research on Cancer (see SM, section 2).

To assess the independent effect of a pollutant on human health and welfare, the EPA's approach considers multiple lines of evidence gathered from various scientific fields, spanning atmospheric physics and chemistry, exposure science, dosimetry, toxicology, statistics, data science, clinical medicine, and epidemiology. The agency systematically identifies, evaluates, and summarizes the relevant peer-reviewed scientific evidence. In this process, the EPA assesses whether there is consistency of effects within a discipline, coherence of effects across disciplines, and evidence of biological plausibility. Thus, the causality determinations developed for an air pollutant and a specific health outcome, such as respiratory effects or mortality, reflect the assessment of the collective body of evidence, rather than a single line of evidence or the use of a single statistical method (*4*). This multidisciplinary framework has been embraced widely by the scientific community as the appropriate process for public health applications (*5*).

Since its inception, this causal framework rooted in the weight of the evidence has been continuously improved through extensive input from CASAC during prior pollutant reviews, involving 11 CASAC panels and 138 individuals (*6*). These improvements have come from building a base in the scientific literature over time, have allowed for newly developed statistical methods to be applied to air pollution studies, and have been supported broadly by CASAC and the scientific community.

The process matters. Under the Clean Air Act's National Ambient Air Quality Standards, the causal determinations developed in the EPA science assessment are used in the risk and exposure assessment and policy assessment to evaluate the impacts of setting air pollution standards at different levels. Together, these three documents are what the EPA administrator will use to set air pollution standards at a level that will protect public health with an adequate margin of safety, as the Clean Air Act requires. Thus, it is crucial that the EPA science assessment reflect the current scientific understanding of a pollutant's effects on health and welfare.

MANIPULATIVE CAUSATION

An alternative framework for determining the linkages between air pollutants and health outcomes in the EPA process is now being promoted by the current CASAC chair, Louis Anthony (Tony) Cox Jr. Rather than look at the weight of the evidence from studies across different fields and different study designs, members of CASAC are proposing in a draft letter that the EPA instead limit the studies that inform its causality determinations to those that can pass a specific narrow approach called manipulative causality (*7*). Cox will oversee the committee's review of the science assessment and related EPA documents and shepherd the development of a scientific recommendation to the EPA administrator on what level of ambient particulate mat-

[1]Center for Science and Democracy, Union of Concerned Scientists, Cambridge, MA, USA. [2]Harvard T. H. Chan School of Public Health, Boston, MA USA. Email: ggoldman@ucsusa.org

Efforts to weaken the process by which the U.S. Environmental Protection Agency obtains independent science advice on the health impacts of air pollution could affect air pollution standards.

PHOTO: TETRA IMAGES/ALAMY STOCK PHOTO

ter will protect public health with an adequate margin of safety.

Under this framework, to justify regulatory action, air pollution epidemiological studies must demonstrate manipulative causation, that is, there must be direct evidence that the implementation of a regulatory action and/or a reduction in pollutant exposure leads to a health benefit. As an attempt to be more precise from a statistical viewpoint, the position argues, in the context of a single epidemiological study, it is necessary to apply causality tests, such as the one implemented by the Causal Analytics Toolkit (CAT), proposed by Cox himself, and/or other existing statistical approaches (Granger causality, information relations in directed acyclic graph models, and Bayesian networks) (see SM section 3). The CASAC chair argues that the majority of current epidemiological studies considered by the EPA only provide evidence of an association (and not evidence of causation) between exposure to air pollution and health effects because, he falsely claims, they do not adjust for confounders (such as weather, demographic, or socioeconomic variables), and therefore, they are not proving manipulative causation.

In principle, attempting to assess causality from observational data in air pollution epidemiology can be viewed as a reasonable framework to address the general issue of confounding bias in individual studies. New statistical methods for the analysis of epidemiological studies on air pollution and health can inform and improve the EPA's approach to its science assessment. Indeed, this is the value of the weight-of-the-evidence approach, which is open to new advances in all fields, including causal inference studies. But instead of allowing these ideas to be introduced, debated, peer reviewed, and advanced in the scientific literature, the CASAC chair suggests that this process be largely skipped and that one specific approach for the analysis of epidemiological data, from a field that is still in its infancy, should trump all other kinds of scientific knowledge.

Further, a requirement of manipulative causation fails to recognize the full depth and robustness of existing approaches in epidemiology, statistics, and causal inference and the degree to which they deal with confounding factors. To study environmental hazards like air pollution, we must rely on analyses of observational data. Randomized control trials are not possible (or ethical) when studying environmental hazards. The great majority of epidemiological studies are designed to estimate how changing an exposure leads to a change in health outcomes while adjusting for confounders, that is, keeping fixed all the other variables that may affect outcomes (such as weather, income, co-pollutants, etc.). Many of the peer-reviewed epidemiological studies included in the EPA's science assessments rely on careful selection of the study design (e.g., time series, prospective cohorts, quasi-experiments), and these studies adjust for confounding bias to infer causality. Many of these studies use regression methods and include the confounders as covariates. Other studies use methods for causal inference and rely on matching, comparing communities or individuals that have different exposures but are matched with respect to the value of the confounders (e.g., individuals with the same education level but different air pollution exposure; see SM section 4). Other studies rely on quasi-randomization (*8*). It has not been convincingly shown that a manipulative causation framework and Cox's proposed tests for causality would be clearly superior to these rigorous and well-vetted approaches.

It remains to be seen whether methods for causal inference such as proposed by CASAC members will become mainstream in air pollution epidemiology. But from the current standpoint, manipulative causation and Cox's causality tests are among many tools in the analytical toolbox. It's not obvious based on current bodies of literature that these new approaches are so powerfully and obviously an indictment of other methods and conclusions, that we should abandon all that we've learned from other approaches, and dismiss all the epidemiological evidence accumulated so far from many disciplines. This all reflects a very normal phenomenon across all science: All methods come with assumptions and have their own strengths and weaknesses, so using varying methods can lead to varying views on a phenomenon.

What matters is the study design and the ability to assess in a transparent way all of the potential sources of confounding bias and error, peer review, and independent reanalyses by experts in the field (*9*). When charged with the task of assessing the weight of evidence of harmful effects from exposure to air pollution, scientists must, and indeed always have, integrate knowledge across many scientific fields and assess all the potential sources of uncertainty. The CASAC chair's proposal suggests skipping this process.

The EPA's mandate under the Clean Air Act requires the agency to protect public health, including within sensitive subgroups (such as children and the elderly), with an adequate margin of safety. To achieve this, the law allows the EPA to be flexible in deciding what an adequate margin of safety is. This is crucial for ensuring that those most sensitive to harm from air pollution are truly protected. The proposed manipulative causation framework and proposed statistical tests of causality, however, place a nearly unattainable burden of proof on the scientific community, and this is unlikely to protect those who need it most.

To be clear, 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. Furthermore, causal inference approaches tend to be more robust to violation of assumptions regarding the form of the statistical model when controlling for confounding bias. There is a literature on methods for casual inference applied to air pollution studies, including the role of causality in data-driven science to inform air pollution regulatory actions (see SM section 5). Regardless, air pollution regulations must be based on existing evidence and demonstrated inference methods that arise from review of existing literature.

In the case of particulate matter, the scientific community has taken several steps to increase the credibility of the results of the epidemiological studies and their ability to infer causality from analyses of observational data. The Health Effects Institute (HEI), a highly regarded independent research institute funded primarily by the EPA and the motor vehicle industry, appointed an independent panel of scientists to reanalyze the results of the landmark American Cancer Society and Harvard Six-Cities Studies that demonstrated the link between long-term particulate exposure and premature death, and indeed they validated the conclusions of the original study (see SM section 6). More recently, the HEI has funded three epidemiological studies to be conducted by three separate teams on three separate populations in three separate locations (United States, Canada, and Europe) to address the defined scientific question of how low-level exposure to fine particulate matter is harmful to human health—a question that arose from the prior particulate matter review completed in 2012. The U.S. team recently published two studies of Medicare data, and used two different study designs (cohort and case cross-over) to estimate the effect of long- and short-term exposure to particulate matter on mortality for the same study population. Both studies found strong evidence of increased risk of mortality at levels well below the safety standards for particulate matter (*10*, *11*).

Despite this cautious, robust, and repeated approach, in its 7 March draft letter, the current CASAC could not reach consensus on the scientific evidence for the link between particulate exposure and mortality. This is perhaps unsurprising given the lack of epidemiological expertise involved. Breaking with historical CASAC member expertise, the EPA leadership declined to place an epidemiologist on the current CASAC. Furthermore, the committee would benefit from the expertise of the particulate matter review panel that the EPA disbanded last October, breaking with decades of precedent of pollutant-specific review panels supplementing the expertise of the seven-member CASAC. The agency also failed to convene a similar review panel for updating the ground-level ozone standard, and the agency expedited review timelines for both pollutants. This means less public input and fewer opportunities for independent scientists, including experts in epidemiology and statistics, to consider and debate this new argument of manipulative causation.

Together, EPA's nixing of the pollutant review panels, the expedited timelines for review of particulate matter and ozone standards, and this narrow view of testing manipulative causation now proposed by the CASAC's chair have proven unpopular among scientists, including experts in the field of causal inference and data science. Three separate letters, penned by 15 members of the dismissed particulate matter review panel, 17 former members of the previous ozone review panel, and 7 previous CASAC members, express the concern about the process and scientific substance of the particulate and ozone reviews led by Cox. Separately, 206 air pollution and public health experts have called on the EPA to reconvene the disbanded particulate matter review panel (see SM section 7).

The particulate matter standard is also being targeted in other ways. The EPA-proposed rule Strengthening Transparency in Regulatory Science carries forward an idea Congress has raised for years but failed to pass (*12*). The rule would restrict the studies that the EPA can use in regulatory decisions by declaring that "the dose response data and models" that underlie regulations must be transparent and accessible to the public. Such restrictions would severely hamstring the EPA's ability to protect people from ambient air pollution. Although some studies do rely on Medicare claims data and therefore would be able to comply with such a requirement, the sweeping proposal raises concerns about study subject privacy regarding medical records, intellectual property, and reproducibility, among other challenges (see SM section 8).

The sum of these changes to the process and scientific approach to setting the particulate matter standard could have far-reaching effects. Avoided particulate pollution accounts for some one-third to one-half of the total monetized benefits of all major federal regulations (not just air quality regulations) (*8*). If the particulate standard is weakened, those benefits would drop in value, and the many public health protections that require cost-benefit analysis to be implemented would be at risk.

Weakening the EPA's long-standing processes for assessing the health impacts of air pollutants could erode the agency's ability to obtain independent science advice on agency decisions on public health protections. In any case, this could ultimately lead to weakening of ambient air pollutant standards. A science assessment that fails to provide a comprehensive look at the relationship between an air pollutant and health effects will yield a subsequent risk and exposure assessment and policy analysis that are flawed, and these crucial documents feed into the EPA administrator's decision on where to set air pollutant standards. Without a robust process to ensure that decision-makers have access to the best available science, policy decisions are unlikely to protect public health.

If the particulate matter and ozone standards are loosened now or in future reviews, people will suffer the consequences. More than 23 million Americans live in areas that exceed the current particulate matter standard, and more than a third of the nation's population lives in areas that exceed the current ozone standard (see SM section 9). If the administration sets air pollution standards that fail to rely on the weight of the evidence on air pollution and health, not only are we casting scientific progress aside, but we risk the health of thousands breathing unhealthy air. As a policy analyst observed, "Science without policy is science, policy without science is gambling" (*13*). ■

REFERENCES AND NOTES

1. J. Tollefson, *Nature* **559**, 316 (2018).
2. https://yosemite.epa.gov/sab/sabproduct.nsf/MeetingCalCASAC/4F40665AD1DDCEF6852583A000645464?OpenDocument
3. G. T. Goldman *et al*., *Science* **355**, 696 (2017).
4. U.S. EPA, Preamble to the Integrated Science Assessments (ISA). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-15/067 (2015).
5. T. A. Glass, S. N. Goodman, M. A. Hernán, J. M. Samet, *Annu. Rev. Public Health* **34**, 61 (2013).
6. J. Vandenberg, Letter to L. A. Cox Jr., 20 February 2019. U.S. Environmental Protection Agency; https://yosemite.epa.gov/sab/sabproduct.nsf//44E735B0EB05DACD852583B500714CA2/$File/JVandenberg+response+to+TCox+ltr+of+121718.pdf.
7. L. A. T. Cox Jr., Clean Air Scientific Advisory Committee (CASAC) Draft Report (03/07/19) to Assist Meeting Deliberations; https://yosemite.epa.gov/sab/sabproduct.nsf/LookupWebProjectsCurrentCASAC/FE50D8FD06EA9B17852583B6006B7499/$File/03-07-19+Draft+CASAC+PM+ISA+Report.pdf.
8. F. Dominici, M. Greenstone, C. R. Sunstein, *Science* **344**, 257 (2014).
9. F. Dominici, C. Zigler, *Am. J. Epidemiol.* **186**, 1303 (2017).
10. Q. Di *et al*., *N. Engl. J. Med.* **376**, 2513 (2017).
11. Q. Di *et al*., *JAMA* **318**, 2446 (2017).
12. www.epa.gov/osa/strengthening-transparency-regulatory-science
13. D. Grey *et al., Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 10.1098/rsta.2012.0406 (2013).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/363/6434/page/suppl/DC1

Published online 21 March 2019
10.1126/science.aaw9460