The potential-outcomes framework is an appealing new approach that imposes a degree of formal conceptual modeling beyond traditional epidemiologic methods for assessing associations between air pollution and health. However, it introduces a number of additional factors to consider when selecting intervention and especially control conditions that call for forward-thinking research designs. We propose that researchers seeking to implement the potential-outcomes framework consider the use of prospective designs that provide more opportunities to establish well-defined intervention and control populations and determine causal relationships between air quality and health. In implementing these prospective research designs, collaboration between researchers and those who implement the interventions can improve the understanding of how a planned intervention actually occurs, thereby improving the characterization of emissions and air quality responses to the intervention. By looking ahead, epidemiologists can take advantage of upcoming regulatory interventions to design successful health outcomes research programs.

accountability; air pollution; health outcomes; regulations

Abbreviation: EPA, Environmental Protection Agency.

Editor’s note: A counterpoint to this article appears on page 1133.

Recent regulations issued by the US Environmental Protection Agency (EPA) that cover power-plant and mobile-source emissions coupled with state and regional programs addressing emissions from goods movement are expected to decrease important contributors to poor air quality, resulting in substantial air quality and health benefits. For example, the US EPA Mercury and Air Toxics Standards are predicted to reduce the number of premature deaths from exposure to particulate matter by as many as 11,000 annually in 2016 (1). The actual health benefits of these stationary and mobile-source programs will depend on specific decisions made by utility companies, rates of turnover in the mobile-source fleet, and effects of any new regulations, economic growth, and energy use, among other factors. Understanding the effectiveness of interventions can serve multiple purposes by providing information that can improve future regulations, by identifying technologies or approaches that can be applied in other areas or to other sectors, and in setting priorities for additional research to improve effectiveness.

In this issue of the Journal, Zigler and Dominici (2) provide an approach for evaluating causal effects of air pollution interventions using a potential-outcomes framework that imposes a degree of formal conceptual modeling beyond traditional epidemiologic methods for assessing associations between air pollution and health. Health outcomes or accountability research related to air pollution has been growing over the past decade in response to research initiatives by the US EPA and nongovernmental organizations, such as the Health Effects Institute and NARSTO (3, 4). However, many earlier studies in this area used relatively traditional methods of epidemiology that focus on establishing associations between changes in air pollution levels and changes in health, in which causality between specific interventions and health outcomes is inferred, not directly assessed.

Zigler and Dominici (2) provide an appealing approach to health outcomes research by focusing on the design of research to elucidate the causal pathway from intervention to health outcome. The potential-outcomes framework they
propose is a useful contribution to the health outcomes liter- ature in its more formal consideration of the nature of the causal relationship between interventions and outcomes. However, at the same time, it is a rather theoretical approach, and it introduces additional factors to consider when selecting intervention and especially control conditions that call for forward-thinking research designs. Implementation of such a framework may not be possible for some types of interventions, and careful evaluation of specific candidate interventions will improve the likelihood for success.

In the present invited commentary, we discuss designs of regulatory outcome studies and make the case that prospective designs for these types of studies provide the best chance for establishing causal relationships between regulatory or other interventions and improvements in air quality and health. We also highlight the need for collaboration between epidemiology researchers and scientists with knowledge of how actions are taken, as well as implementers of the interventions who can best provide data on how a planned intervention actually occurs (e.g., what are the actual emissions controls applied and how much resulting emissions reduction actually occurs). Finally, we highlight upcoming opportunities for exercising the potential outcomes framework and urge epidemiologists to focus on those upcoming regulatory interventions that might be more amenable to health outcomes research.

PROSPECTIVE DESIGN OF HEALTH OUTCOME STUDIES

With few exceptions, health outcomes research has focused on retrospective opportunistic evaluations of interventions, taking advantage of existing observational air quality data (often collected for regulatory purposes) and administrative health data (hospitalization and mortality rates). A few recent studies, such as those conducted regarding air quality changes during the Beijing Olympics (5), were designed prospectively. Retrospective studies suffer from irreducible uncertainties introduced by spatial and temporal confounders, even if they make use of the best available data on emissions, air quality, and health to analyze the effects of air quality improvements that result from interventions.

In contrast, prospectively designed quasi-experimental studies can take better advantage of upcoming natural experiments that result from regulatory or other interventions by carefully collecting data on air pollution mixtures, population health outcomes, and potential confounding factors in both locations expected to be affected by the intervention and control locations with similar populations that are not expected to be affected by the intervention. Data should be consistently collected before, during, and after the intervention.

USING BOTH TIME AND SPACE TO DEFINE INTERVENTION AND CONTROL POPULATIONS

Assessing health effects of interventions is made difficult by the temporality of an action and of its effects on air quality, as well as by the possibility that other confounders can interfere with effects of the intervention as time after the intervention increases. Selecting a control population that is similar to the intervention population in the temporal pattern of potential confounders but that has relatively stable air quality can help to isolate the causal relationship by more clearly establishing the counterfactual.

Zigler and Dominici (2) discuss the use of controls in time or space. One key requirement is that temporal trends (in air pollution, socioeconomic status, and health) in the control areas be comparable to those in the intervention location. Control populations from locations that are relatively close to the intervened area may be preferable to populations from more distant areas because there is a higher chance of comparability. At the same time, areas closer to the intervened areas may be affected by the intervention, and therefore using those populations as controls may lead to overcontrolling for and underestimation of policy-related effects. Selection of appropriate control locations is especially difficult when evaluating broad-scale regulatory measures that potentially affect air pollution at regional or national scales, such as the Clean Air Act Title IV regional trading program for power-generating facilities, the example chosen by Zigler and Dominici.

Peel et al. (6) evaluated effects of the Atlanta Olympics (and associated traffic management programs) on air quality and emergency department visits and found that although air quality did improve in Atlanta during the Olympics, it also improved throughout the southeastern United States, likely because of meteorological conditions that were less conducive to ozone formation during that period. Without careful consideration of air quality trends in locations that were not expected to be affected by the Olympics, the investigators might have concluded that the traffic management programs used during the Olympics caused the reductions in ozone levels.

It is important to identify those variables that ideally should be collected at baseline (before the intervention) and tracked over time to control for potential confounding by secular trends (e.g., greater use of statin drugs to control cardiovascular risk factors) because they can lead to biased assessments of the health outcomes of air quality improvements. Much can be learned about these types of variables from existing retrospective health outcomes studies. For example, in the recent follow-up study by Dockery et al. (7) of health outcomes after the Dublin coal ban, the investigators discovered that earlier findings that the coal ban decreased cardiovascular mortality rates were not confirmed once they included a spatially isolated comparison population as a control group. They determined that the previous conclusion that the coal ban in Dublin was associated with reductions in cardiovascular outcomes was likely overstated because those reductions were potentially associated with secular trends in cardiovascular mortality that were unrelated to the coal ban.

One key factor in prospective planning of intervention research is that the same types of data need to be collected for both the control and intervention population groups. In retrospective study designs, in most cases, more data are available for intervention populations (which tend to be from locations that have air quality problems and are thus more likely to have comprehensive monitoring data) than for control populations. Ideally, the same types of information would be available on air quality, health conditions, the economy, and other factors.
correlated with health outcomes that change over time so that it would be possible to control for factors that can confound the analysis. However, air quality monitors are not uniformly located, and this limits the ability to select populations not affected by national-scale air quality interventions. With prospective planning, it may be possible to put into place special purpose monitors coupled with additional information from satellites or air quality modeling to track air quality in those control locations. In addition, it may be possible to collect additional information on health outcomes for selected at-risk subpopulations to add additional power to the experimental design, as was done in the Beijing Olympics study (5). This requires careful thinking about data collection in research design. Without such data, researchers are left trying to statistically control for potential confounders, hoping to tease out the causal signal.

INVOKE THE “INTERVENERS” IN THE RESEARCH DESIGN PROCESS

It is important to carefully consider potential differences between intended effects of an intervention and those of the actual implementation. For example, Kelly et al. (8) found that nitrogen dioxide concentrations actually increased in the London congestion charging zone, likely as the result of the parallel introduction of diesel particle traps on buses, which had higher direct nitrogen dioxide emissions. Peel et al. (6) consulted with traffic planners and determined that traffic volumes during the Atlanta Olympics had shifted in time but had not actually declined overall. Without this thorough understanding of the set of interventions occurring during the time period, investigators could be left with very confusing results. Health outcomes researchers will benefit from involving regulators and others involved in implementing air quality interventions to understand how interventions actually occur; for example, did the intervention result in changed behavior, emissions reductions, air quality improvements, and changed health outcomes as expected?

This is clearly an important difference between quasi-experimental approaches and traditional controlled experiments because the researcher does not control the intervention, and thus it is important to confirm that an actual reduction in the level of air pollution was achieved and whether there were any other interventions that occurred at the same time. In addition, the assignment of populations to the intervention and control groups is not random. This highlights the need to make sure that the intervention and control groups are as similar as possible in important variables that may confound the air pollution–health signal. In many cases, what might seem to be a good candidate intervention for a health outcomes study might actually be very difficult because of the complexities of defining the intervention and control populations.

For example, evaluation of the impact of nonattainment designations for national ambient air quality standards seems at first to be a good candidate, because the US EPA clearly identifies counties as in attainment or not in attainment with air quality standards. However, designations for nonattainment are problematic because it is not clear which group should be defined as the intervention group—that is, counties are included in nonattainment areas either if they have an air quality monitor that exceeds the national ambient air quality standards or if they are determined to have emissions that are contributing to a nearby air quality monitor that is not in attainment. In addition, downwind counties may also receive air quality benefits from emissions reductions in non-attainment counties and thus may not be good controls.

PLANNING AHEAD FOR PROSPECTIVE OUTCOMES STUDIES

Demonstrating the effectiveness of potential-outcomes thinking will benefit from careful evaluation of types of interventions that would be best suited to testing different quasi-experimental designs and assessing the need to collect data on intervention and control groups that would allow for the use of such designs. The biggest challenge in these and other approaches is defining appropriate control groups.

There are several large interventions, both regulatory and nonregulatory, that are planned over the next decade. These include implementation of the US EPA Mercury and Air Toxics Standards, which are expected to decrease emissions of sulfur dioxide and several hazardous air pollutants, several mobile-source regulations on fuels and engines that are expected to decrease volatile organic compounds and nitrogen oxides emissions, and the expansion of the Panama Canal, which is expected to increase goods shipments through eastern US ports and potentially affect emissions related to both ports and goods-movement activities. Evaluation of each of these interventions before they occur to determine which might present the best opportunities for air quality and health outcomes research may best allow for the design of quasi-experiments with an appropriate control group and give the best opportunity to establish a causal link between interventions and health effects.

ACKNOWLEDGMENTS

Author affiliations: Health and Environmental Impacts Division, Office of Air Quality Planning and Standards, US Environmental Protection Agency, Research Triangle Park, North Carolina (Bryan Hubbell); and Health Effects Institute, Boston, Massachusetts (Daniel Greenbaum).

The views expressed in this commentary are those of the authors and do not represent the official views of the US Environmental Protection Agency.

Conflict of interest: none declared.

REFERENCES


