NATIONAL ACADEMIES Sciences Engineering Medicine

NATIONAL ACADEMIES PRESS
Washington, DC

Advancing the Framework for Assessing Causality of Health and Welfare Effects to Inform National Ambient Air Quality Standard Reviews

https://www.nationalacademies.org/publications/26612

Committee on Assessing Causality from a Multidisciplinary Evidence Base for National Ambient Air Quality Standards

Board on Environmental Studies and Toxicology

Division on Earth and Life Studies

Board on Mathematical Sciences and Analytics

Division on Engineering and Physical Sciences

Consensus Study Report

high-level conceptual definition of causality. The committee does not mean to imply that individual studies that do not use a potential outcome framework should not be considered in the ISA process. The field of causal inference is rapidly evolving, and a variety of causal inference approaches can be used in the context of individual studies and provide insights relevant for the ISA causality determinations.

DEFINITION OF "CAUSE" FROM THE PREAMBLE

The Preamble (EPA, 2015a) uses a definition of causality articulated originally in the 1964 Surgeon General's report (HEW, 1964) that is highly related to the idea of potential outcomes: "The 1964 Surgeon General's report on tobacco smoking defined 'cause' as a 'significant, effectual relationship between an agent and an associated disorder or disease in the host'" (EPA, 2015a, p. 18). However, determining an effectual relationship between an agent and associated disorder is complicated by a lack of direct observation. As such, EPA integrates and synthesizes evidence from a wide and comprehensive body of scientific literature. The Preamble summarizes EPA's approach to evaluating causal relationships from this synthesized evidence as follows:

> In its evaluation and integration of the scientific evidence on health or welfare effects of criteria pollutants, the U.S. EPA determines the weight of evidence in support of causation and characterizes the strength of any resulting causal classification. The U.S. EPA also evaluates the quantitative evidence and draws scientific conclusions, to the extent possible, regarding the concentration-response relationships and the loads to ecosystems, exposures, doses or concentrations, exposure duration, and pattern of exposures at which effects are observed. (EPA, 2015a)

EPA relies on the Bradford Hill aspects of association (Hill, 1965) to guide its weight of evidence approach. Those aspects are aligned with a potential outcomes/counterfactual way of thinking about causation used in epidemiological studies, in terms of relating different exposures (i.e., causes) to outcomes (i.e., effects). The potential outcome framework is useful for assessing the evidence of causality provided by an individual study. The Bradford Hill aspects of association are useful for then integrating evidence of causality from multiple sources, taking into account the quality and relevance of the evidence from each of the individual studies.

EVALUATING CAUSALITY THROUGH MECHANISTIC UNDERSTANDING

The above description defining causal effects does not require any knowledge of why such effects occur. An alternative approach to causality is known as "mechanistic causality." Mechanistic causality requires an understanding of the causal mechanism—that is, all the steps underlying an observed association need to be explicit and have a genuine explanation (Campaner, 2011). To achieve mechanistic knowledge is to be able to determine not only that some factors contribute to some effects, but also how they contribute, revealing the continuous and dynamic processes linking causes with their effects. A mechanistic perspective on causality holds that a causal relationship can be revealed through identification of the specific mechanisms that describe how a given cause affects the study outcome. More detailed definitions and discussion of variations of this approach can be found in Salmon (1984), Glennan (1996), and Machamer et al. (2000). Writing from a classical physics perspective, Glennan (1996) defined mechanisms as "complex systems whose 'internal' parts interact to produce a system's 'external behavior,'" arguing that "events are causally related when there is a mechanism that connects them."

A major part of establishing causality includes some aspects of mechanistic understanding. Knowledge of underlying mechanisms, however imperfect, can help clarify which apparently causal relationships are plausible and potentially relevant. For example, mechanistic understanding can help inform how causal effects should be defined (e.g., the appropriate time lag between the exposure of interest and the outcome of interest), and how well one can learn about the causal relationships of interest (e.g., to help understand how plausible the underlying assumptions of studies are, such as whether confounding is sufficiently dealt with). Perfect knowledge of all mechanisms involved in all the effects of air pollution on public health or welfare is currently (and may always be) unattainable.

For example, Brook and colleagues (Brook et al., 2010) have elucidated many pathological mechanisms that lend biological plausibility to the adverse effects of fine particulate matter on cardiovascular health. In this context, mechanistic causality could be determined if the pathways and biological mechanism of how air pollutants lead to cardiovascular diseases are fully identified. An understanding of these mechanisms can also help identify, for example, which confounders are important to adjust for in studies relating fine particulate matter to cardiovascular health. Another example of a mechanistic approach is provided in Box 3.1 relating particulate matter to visibility.

CAUSAL MODELS

A concept that brings together the ideas of mechanistic causality and the definition of causal effects as defined using potential outcomes are "causal models." Causal models—which aim to establish an understanding of causal effects in part through the clear articulation of causal assumptions—must be distinguished from purely statistical models. Given the unobserved potential outcomes and "fundamental challenge of causal inference" described above, assumptions or knowledge encoded in a causal model often involve the potential outcomes. These assumptions may or may not have empirical implications or be empirically testable. Statistical models involve modeling observed data only, and all assumptions in statistical models are empirically testable. For example, temporal ordering (i.e., the ordering of events in time) is a type of knowledge encoded in a causal model. This means that in a study assessing the causal effect of A on Y, it is necessary to know a priori whether A is antecedent to Y, as the cause must precede their effects.

Another example of knowledge encoded in a causal model assumption is randomization. Randomization works because the randomized exposure is independent of the potential outcomes,

BOX 3.1
A Mechanistic Approach to Causality: Relating Particulate Matter to Visibility

An example of the mechanistic approach applied to causality in past NAAQS processes is that which related the effects of particulate matter on visibility (Malm, 2016). The physical mechanisms by which air pollutants affect the optical characteristics of the atmosphere include the scattering and absorption of light by gases and particles, with scattering by particles being the predominant mechanism affecting visibility at most times and locations (Malm, 2016; Middleton, 1957). The efficiency of particle light scattering is greatest as the particle size approaches the wavelengths of visible light (0.4 to 0.7 microns), and so small particles ($PM_{2.5}$) are typically the largest contributors to visibility impairment (Malm, 2016; van de Hulst, 1958). Scattering of light by small particles is referred to as Mie scattering, after the German Physicist Gustav Mie, who developed calculations of light scattering by particles with sizes like the wavelengths of visible light, as a function of wavelength, particle size, and index of refraction (Mie, 1908). Mie's calculations are based on direct solutions to Maxwell's equations (Maxwell, 1865) for electromagnetic radiation, thus linking the most important mechanism affecting visibility impairment directly to fundamental laws of physics. Further details and references are provided in the 2019 PM ISA (EPA 2019c, Section 13.2), and the 2009 PM ISA (EPA, 2009, Section 9.2).

by design. When assessing the causal effect of a randomized exposure A on an outcome Y, knowledge of such independence must come from sources external to observations on A and Y, namely knowledge of the study design—that randomization to exposure conditions occurred. This knowledge of randomization will have implications in the data—in particular similarity of covariates distributions across exposure groups, on average—but data analysis will not guarantee or prove that randomization occurred. Another example of an untestable causal assumption that is often encoded or assessed using a causal model is whether a given set of pre-treatment variables is sufficient to control for confounding (Greenland et al., 1999; Rubin, 1978). Causal models are informed by a priori understanding of the underlying mechanisms, and make explicit any assumptions that are required to identify causal effects from observed data (as discussed above, a causal effect is identified if it can be uniquely computed from the distribution of observed variables). In brief, statistical models allow assessment of whether two variables are related in a dataset. A causal model articulates the assumptions—and possibly their plausibility—required to interpret that relationship as causal (Carone et al., 2020).

In this report, the committee adopts the term causal model to refer to this set of a priori assumptions about unobserved variables, including the unobserved potential outcomes and how they relate to observed variables. The committee also distinguishes between causal models and statistical

BOX 3.2
Useful Definitions

Causal effect. A quantity defined in terms of a causal model. Estimating a causal effect requires the assumptions in the causal model to be correct and involves consideration of the unobserved (and unobservable) counterfactual outcomes (Pearl, 2009). Causal effects cannot be learned from a statistical model and data alone. Causal effects are also sometimes referred to as "causal parameters."

Causal inference. The process of estimating, after having postulated a causal model, the effect of a cause on an outcome (health or welfare) from a randomized or observational study (Pearl, 2009). Common approaches include counterfactual causal inference, manipulative causal inference, and mechanistic causal inference (their definitions are included in this chapter).

Causal mechanism. The processes or pathways through which an outcome is brought into being.

Causal model. A model, usually represented mathematically, incorporating a priori causal assumptions about the relationships within an individual system or population. Causal models may and often posit assumptions about unobserved variables, such as exogenous factors or potential outcomes.

Confounding. The effect of a variable (the confounder) that influences both the dependent variable (e.g., health or welfare outcome) and independent variable (e.g., exposure to air pollution), resulting in a non-causal association in the data. An example involves three phenomena: carrying matches, smoking, and lung cancer. The causal effect of carrying matches on lung cancer is zero, but the statistical association may be positive in any given dataset. This artifact (from a causal point of view), which arises because smokers are likely to both carry matches and develop lung cancer, is an example of confounding. While a confounder may be associated with both the dependent (e.g., health outcome) and independent (e.g., exposure) variable, it should not also fall on the causal pathway.