ERIC Notebook

Second Edition

Causality

Second Edition Authors:

Lorraine K. Alexander, DrPH

Brettania Lopes, MPH

Kristen Ricchetti-Masterson, MSPH

Karin B. Yeatts, PhD, MS

The primary goal of the epidemiologist is to identify those factors that have a causal impact on development or prevention of a health outcome, thereby providing a target for prevention and intervention. At first glance, causality may appear to be a relatively simple concept to define. However, adequately distinguishing causal agents from non-causal agents is not an easy task from an epidemiologic perspective. Unfortunately, there is no elementary parameter that can be measured to provide a definitive answer when determining causality. Rather, there are a series of criteria that have been developed and refined over the years that now serve as the guideline for causal inference. The most important point to remember is that causality is not determined by any one factor, rather it is a conclusion built on the preponderance of the evidence.

Epidemiologist Austin Bradford Hill is credited with identifying the nine factors that constitute the current standard for determining causality (1965). In his article, Hill expanded upon criteria that had previously been set forth in the report *Smoking and Health* (1964) by the United States Surgeon General. Below is a discussion of the nine criteria defined by Hill to be utilized in the determination of causality.

It is important to note that satisfying these criteria may lend support for causality, but failing to meet some criteria does not necessarily provide evidence against causality, either. Hill's causal criteria should be viewed as guidelines, not as a "checklist" that must be satisfied for a causal relationship to exist.

Hill's causal criteria

Strength of association

Strength of association between the exposure of interest and the outcome is most commonly measured via risk ratios, rate ratios, or odds ratios. Hill believed that causal relationships were more likely to demonstrate strong associations than were non-causal agents. Smoking and lung cancer is a perfect example where risk ratios, rate ratios, and odds ratios are in the 20 to 40 range when comparing smokers to non-smokers. However, weak associations as demonstrated by the risk ratio, rate ratio, or odds ratio should not be taken as an indication of non-causality. This is particularly true when the outcome of interest is common.

An example of a common outcome that exhibits a weak association to smoking is cardiovascular disease (CVD). Yet even with a weak association, evidence supports the casual nature between smoking and the development of CVD. Furthermore, one should not assume that a strong association alone is indicative of causality, as the presence of strong confounding may erroneously lead to a strong causal association.

Consistency of data

This tenant refers to the reproducibility of results in various populations and situations.

Consistency is generally utilized to rule out other explanations for the development of a given outcome. It should also be noted that a lack of consistency does not negate a causal association as some causal agents are causal only in the presence of other co-factors. In general, the greater the consistency, the more likely a causal association.

Specificity

This criterion has been proven to be invalid in a number of instances, with smoking being the primary example. Evidence clearly demonstrates that smoking does not lead solely to lung carcinogenesis but to a myriad of other clinical disorders ranging from emphysema to heart disease. On the other hand, there are certain situations where a 1 to 1 relationship exists, such as with certain pathogens which are necessary to produce a specific disease. Tuberculosis is a good example.

Temporality

This criterion has been identified as being the most likely to be the sine qua non for causality, i.e. it is absolutely essential. For an agent to be causal, its presence must precede the development of the outcome. Lack of temporality rules out causality. An example found in the literature is the relationship between atrial fibrillation (AF) and pulmonary embolism. Current wisdom supports that pulmonary embolism causes atrial fibrillation, however more recent evidence and plausible biological hypothesis suggest that the reverse could be true. Determining the proper course of care may hinge upon discovering if pulmonary emboli can indeed precede and thus perhaps cause the development of atrial fibrillation.

Dose-response

The presence of a dose-response relationship between an exposure and outcome provides good evidence for a causal relationship; however, its absence should not be taken as evidence against such a relationship. Some diseases or health outcomes do not display a dose-response relationship with a causal exposure. They may demonstrate a threshold association where a given level of exposure is required for disease or health outcome initiation, and any additional exposure does not affect the outcome.

Biological plausibility

Support for this criterion is generally garnered in the basic science laboratory. It is not unusual for epidemiological conclusions to be reached in the absence of evidence from the laboratory, particularly in situations where the epidemiological results are the first evidence of a relationship between an exposure and an outcome. However, one can further support a causal relationship with the addition of a reasonable biological mode of action, even though basic science data may not yet be available.

Coherence

This term represents the idea that, for a causal association to be supported, any new data should not be in opposition to the current evidence, that is, providing evidence against causality. However, one should be cautious in making definite conclusions regarding causation, since it is possible that conflicting information is incorrect or highly biased.

Experimental evidence

Today's understanding of Hill's criteria of experimental evidence results from many areas: the laboratory, epidemiological studies, and preventive and clinical trials. Ideally, epidemiologists would like experimental evidence obtained from a well-controlled study, specifically randomized trials. These types of studies can support causality by demonstrating that "altering the cause alters the effect".

Analogy

This is perhaps one of the weaker of the criteria in that analogy is speculative in nature and is dependent upon the subjective opinion of the researcher. An example of an analogy is that while infection may cause a fever, not all fevers are due to infection. Absence of analogies should not be taken as evidence against causation.

Other considerations

In addition to assessing the components of Hill's list, it is also critically important to have a thorough understanding of the literature to determine if any other plausible explanations have been considered and tested previously.

Additional models for causality

In additional to Hill's guidelines for causality, several recent models for understanding causality have been developed. These include Kenneth Rothman's component cause theory, counterfactual models, and directed acyclic graphs. We provide brief descriptions below but refer you to the suggested readings section for more information.

Component causes

Kenneth Rothman described the circumstances leading to a health outcome as being parts of one pie chart, or a "causal pie." Without each component in place, the disease or health outcome would not have occurred at that specific point in time.

Example

If you were to slip on an icy sidewalk and break your wrist, there may be a number of factors that contributed to that outcome. First, the fall directly leads to a broken wrist, and this was a necessary component of the outcome. However, which factors led to you walking on the ice and having a fall in the first place, and what other factors influenced your broken wrist? Maybe you were wearing poor footwear and were tugged by your dog who was chasing a squirrel. Maybe you didn't receive enough calcium in your diet, developed osteoporosis, and the fall would not have broken your wrist without weakened bones. Each of these components ultimately led to the outcome.

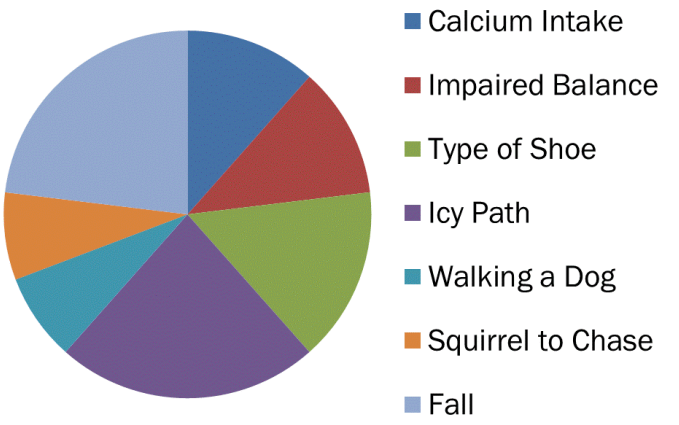

Figure: Causal Pie

Rothman's component causes theory is one way to consider all factors involved in the development of a health outcome.

Counterfactual models

Many statistical models that are used to adjust for confounding are based on counterfactual thinking. These models are based on comparing an exposed group of people to a fictional group of people who are exactly the same except they are unexposed to the key variable. These models try to answer the question: "If this one experience or exposure in the past did not happen to an individual, how would it impact that person's health outcome today?" Of course, this is an impossible situation; we cannot go back in time and change an individual's exposure status and track both outcomes over time. However, using two very similar populations of people – one group that is exposed and another that is unexposed – it is possible to estimate that very counterfactual situation on a group level. Many statistical modeling programs that adjust for potential confounders are modeling a counterfactual scenario to produce a less biased measure of association.

Directed acyclic graphs (DAGs)

DAGs are one method used to create a conceptual diagram that maps the relationships between the main exposure, outcome of interest, and all potential confounders for a given study. Through a DAG analysis, specific rules are followed to determine if confounding might be present for a given research question. Once confounders have been identified and adjusted for, a less biased measure of association can be obtained.

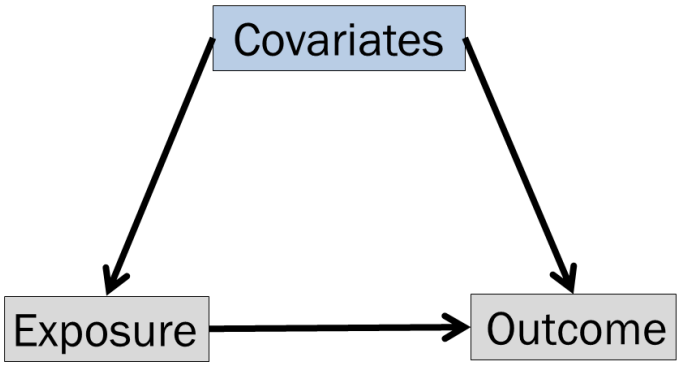

Figure: Basic Directed acyclic graph (DAG)

References and Suggested Readings

Flegel KM. When atrial fibrillation occurs with pulmonary embolism, is it the chicken or the egg? *CMAJ*. 160 (8):1181-2, 1999.

Greenland S, Holland PW, Mantel N, Wickramaratne PJ and Holford TR. Confounding in Epidemiologic Studies. *Biometrics*. 45(4):1309-1322

Hill AB. The environment and disease: association or causation? *Proc R Soc Med.* 58:295-300, 1965.

Mengersen KL. Merrilees MJ. Tweedie RL. Environmental tobacco smoke and ischaemic heart disease: a case study in applying causal criteria. *Int Arch of Occup & Env Health*. 72 Suppl:R1-40, 1999.

Ridgway D. The logic of causation and the risk of paralytic poliomyelitis for an American child. *Epidemiology & Infection*. 124(1):113-20, 2000.

Rothman KJ. Causes. *Am J Epidemiol* 104:587–92, 1976.

Rothman KJ, Greenland S, Poole C, and Lash TL. Causation and causal inference. In: *Modern Epidemiology (3e)*. Edited by: Rothman KJ, Greenland S, and Lash TL. Philadelphia: Lippencott-Raven Publishers; 2008:5-31.

Shrier I and Platt RW. Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology* 2008, 8:70.

United States Department of Health, Education, and Welfare. *Smoking and health:* Report of the Advisory Committee to the Surgeon General of the Public Health Service. Washington, D.C Government Printing Office, 1964 PHS Publ. No. 1103.

Weed DL. On the use of causal criteria. *Int J of Epidemiology*. 26(6):1137-41, 1997.

Dr. Carl M. Shy, Epidemiology 160/600 Introduction to Epidemiology for Public Health course lectures, 1994-2001, The University of North Carolina at Chapel Hill, Department of Epidemiology

Acknowledgement

The authors of the Second Edition of the ERIC Notebook would like to acknowledge the authors of the ERIC Notebook, First Edition: Michel Ibrahim, MD, PhD, Lorraine Alexander, DrPH, Carl Shy, MD, DrPH, Sandra Demming, MPH, Department of Epidemiology at the University of North Carolina at Chapel Hill. The First Edition of the ERIC Notebook was produced by the Educational Arm of the Epidemiologic Research and Information Center at Durham, NC. The funding for the ERIC Notebook First Edition was provided by the Department of Veterans Affairs (DVA), Veterans Health Administration (VHA), Cooperative Studies Program (CSP) to promote the strategic growth of the epidemiologic capacity of the DVA.