ONLINE MATERIAL THAT ACCOMPANIES CHAPTER 7

Box 1. Conceptualizing cause in epidemiology and population health through history

The concept of a cause and the criteria needed in population health studies to establish a causal relation have gone through substantial changes throughout the history of epidemiology. Knowing the historical evolution of how epidemiologists have thought about the causes of poor health sheds light on epidemiology's evolution as a discipline. The various eras of causal thinking align well with the eras of public health (Susser and Susser 1996), and with the dominant public health conditions that were of concern to public health scientists during various time periods. We highlight two paradigms below, as they are still relevant to the work that we do as epidemiologists today.

The Henle-Koch Postulates (Evans 1976)

In the late 19th century, the prevalent diseases of public health concern were predominately communicable diseases such as cholera, typhoid, and tuberculosis. Robert Koch, a German physician and scientist, developed four postulates intended to be generalizable across outcomes in order to establish a causal link between an exposure (in Koch's case, a microorganism) and a particular disease. Along with Friedrich Loeffler, Koch originally articulated three postulates in 1884, then revised them several times over the next several decades as scientific advances led to a more comprehensive understanding of communicable and infectious disease.

Koch's postulates essentially described the following conditions:

1. The microorganism must be present in all organisms that have the disease and must not be present in organisms that do not have the disease.
2. The microorganism must be able to be isolated from the host and grown in pure culture.
3. When the cultured organism is introduced into an otherwise healthy organism, the healthy organism should develop disease.

A fourth postulate specifying that a disease-causing microorganism must be able to be re-isolated from the experimental host and identified as identical to the original was also added to the original three postulates. However, as science progressed, Koch noted that these postulates are not generalizable to all infectious disease or beyond infectious disease. For example, there are asymptomatic carriers of cholera, thus violating the first principle. Nonetheless, they present a straightforward and succinct set of conditions for conceptualizing a cause in the case of communicable illness.

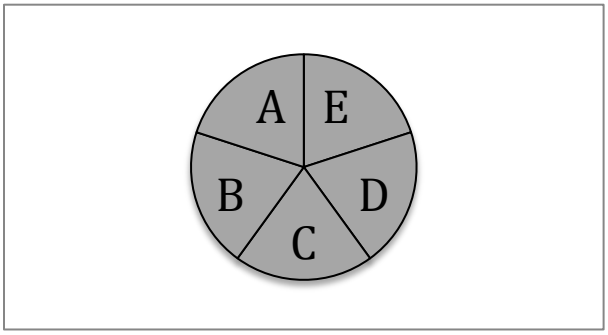
Hill's Causal Criteria (Hill 1965)

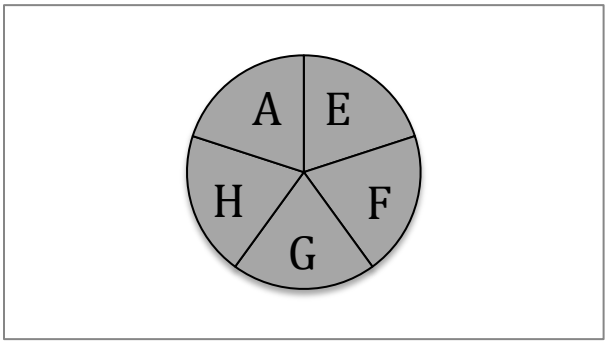
In 1965 Sir Austin Bradford Hill wrote a paper that articulated conditions that he thought would suggest that some factors are more likely to be causes than others. These factors include strength and specificity of the association observed, consistency of observations across studies, temporality (the cause was observed before the health effect), biological gradient (greater exposures are observed to lead to greater effects), plausibility (the existence of a plausible biological mechanism to observe the documented effect), coherence (similar findings in epidemiologic and laboratory findings of particular effects), experiment (randomized controlled trial shows an effect of the exposure), and analogy (where the effect of similar factors is similar).

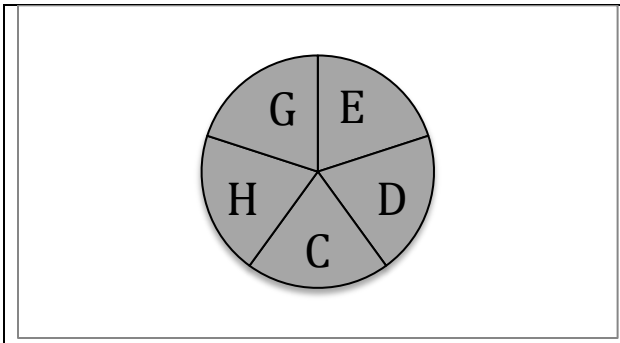
Bradford Hill did not intend his conditions to serve as a check list to determine cause; nonetheless, for many decades, they did serve as signposts against which many factors were weighed when debating whether or not they were causes. Now we know that all but one of the Bradford Hill criteria can be refuted; that is, for example, although observations drawn from experimental studies may indeed suggest that a particular factor is a cause, there are many other causes that we accept without recourse to experiment. The one exception is temporality, which remains a *si ne qua non* for causation. The cause must precede the disease in all circumstances, logically and pragmatically.

Rothman's Sufficient-Component Cause Model (Rothman 1976, Rothman and Greenland 2005)

Kenneth Rothman's sufficient-component cause model has been a mainstay of introductory and advanced epidemiologic methods, and it was the model that we used as a basis for our marble analogy. The sufficient-component cause model uses pies to represent the sufficient causes of disease. Each piece of the pie is a component cause, which means it is necessary in order for the disease to occur but insufficient if there is more than one component cause in the pie. An example is given below:







For a hypothetical disease there are three sufficient pies. Once an individual acquires all pieces in a given pie, disease inevitably occurs. The same concepts of necessity and sufficiency introduced in this chapter apply to the Rothman pies. Piece "E" for example, is in all three sufficient pies and is thus necessary for disease to occur. If we prevent exposure to "E", no disease will occur. Most component causes are neither necessary nor sufficient at a population level, i.e., they are in one or more sufficient causal pies but not all, and they require the presence of other slices in order for the health outcome to occur.

Citations

Evans, A. S. (1976). "Causation and disease: the Henle-Koch postulates revisited." Yale J Biol Med 49(2): 175-195.

Hill, A. B. (1965). "The Environment and Disease: Association or Causation?" Proc R Soc Med 58: 295-300.

Rothman, K. J. and S. Greenland (2005). "Causation and causal inference in epidemiology." Am J Public Health 95 Suppl 1: S144-150.

Box 2. Diethylstillbestrol (DES) and vaginal cancer

The tragic case of diethylstilbestrol (DES) and vaginal cancer illustrates how individuals can be exposed to a disease-causing factor at one point in life with the result of the exposure manifesting much later in the life course. Approved for use in 1938, DES is a synthetic estrogen widely prescribed to millions of pregnant women in the mid 20th century. It was thought that increased estrogen would reduce the risk of miscarriage, premature delivery, and stillbirth. Research in the 1950s confirmed that DES did not actually have an effect on pregnancy outcomes. However, DES was still prescribed until 1971, when clusters of very young women with vaginal cancer (an exceedingly rare condition especially in adolescence and young adulthood) began being reported in the scientific literature. It was discovered that these women were almost all offspring of women who had taken DES during pregnancy (Herbst, Ulfelder et al. 1971). *In utero* exposure to DES is now confirmed to increase the risk for early onset vaginal cancer. Further, daughters of DES-exposed mothers are at increased risk for infertility and adverse reproductive outcomes in their own offspring (Hoover, Hyer et al. 2011), as well as breast and ovarian cancer (Colborn, vom Saal et al. 1993). Sons of DES-exposed mothers are at increased risk for adverse testicular outcomes such as non-cancerous epididymal cysts (Gill, Schumacher et al. 1979). Investigators are now examining health effects in the

grandchildren of DES-exposed mothers. Thus, the causes of adverse health outcomes can begin *in utero* and perhaps before conception even occurs. The dose, timing, and additional exposures through the life course shape the health outcomes observed among those exposed to adverse influences early in life, yet the importance of the early life environment in child, adolescent, and adult health cannot be overlooked.

Citations

Colborn, T., F. S. vom Saal and A. M. Soto (1993). "Developmental effects of endocrine-disrupting chemicals in wildlife and humans." Environ Health Perspect 101(5): 378-384.

Gill, W. B., G. F. Schumacher, M. Bibbo, F. H. Straus, 2nd and H. W. Schoenberg (1979). "Association of diethylstilbestrol exposure in utero with cryptorchidism, testicular hypoplasia and semen abnormalities." J Urol 122(1): 36-39.

Herbst, A. L., H. Ulfelder and D. C. Poskanzer (1971). "Adenocarcinoma of the vagina. Association of maternal stilbestrol therapy with tumor appearance in young women." N Engl J Med 284(15): 878-881.

Hoover, R. N., M. Hyer, R. M. Pfeiffer, E. Adam, B. Bond, A. L. Cheville, T. Colton, P. Hartge, E. E. Hatch, A. L. Herbst, B. Y. Karlan, R. Kaufman, K. L. Noller, J. R. Palmer, S. J. Robboy, R. C. Saal, W. Strohsnitter, L. Titus-Ernstoff and R. Troisi (2011). "Adverse health outcomes in women exposed in utero to diethylstilbestrol." N Engl J Med 365(14): 1304-1314.

Box 3. From deterministic to probabilistic: an example

In Chapter 7 we provided a simple example of a disease that was caused by the interaction of three component causes: X, Y, and Z. We showed how the probability of disease given exposure to X was 26%, and that this probability was less than 100% because we did not simultaneously take account of Y and Z. Now, let us look at a more complicated example.

Suppose there are two ways to get disease: exposure to X, Y, and Z, or exposure to A, B, and C. In our marble analogy, we would say that either of two jars is sufficient to cause disease: a jar containing X, Y, and Z, or a jar containing A, B, and C. Thus, 100% of people exposed to A, B, and C will get the disease and 100% of people exposed to X, Y, and Z will get the disease. Now again imagine that we have a sample of 100 people. In our population, 50 individuals are exposed to X and 50 individuals are not exposed to X. In Table 1 below, we show all combinations of potential exposure to X, Y, Z, A, B, and C that would result in disease, and the number of people in the population with each combination of component causes. For example, there is 1 individual that exposed to A, B, C, X, and Y. This individual will have the disease. Note that this individual will have the disease even if not exposed to X, because this individual is exposed to A,B,C. Therefore exposure to X did not cause disease for this individual because it does not meet the counterfactual definition - this individual would have gotten the disease whether exposed to X or not, because this individual was exposed to A,B, and C. There are two individuals exposed to A,B,C,X, and Z, and both of these individual will be diseased (again, because of their exposure to A,B,C). A total of 25 individuals have a set of component causes that produce disease (either X,Y,Z or A,B,C, or both), leaving 75 individuals in the population disease free.

Box 3, Table 1. Hypothetical example where A, B, C, X, Y, and Z are component causes and all measured

A	B	C	X	Y	Z	Number of people with this combination of component causes	Probability of disease
☑	☑	☑	☑			0	100%
☑	☑	☑	☑	☑		1	100%
☑	☑	☑	☑		☑	2	100%
☑	☑	☑	☑	☑	☑	1	100%
☑			☑	☑	☑	3	100%
	☑		☑	☑	☑	2	100%
		☑	☑	☑	☑	4	100%
☑	☑		☑	☑	☑	1	100%
	☑	☑	☑	☑	☑	1	100%
☑		☑	☑	☑	☑	0	100%
			☑	☑	☑	2	100%
☑	☑	☑	☑	☑	☑	2	100%
☑	☑	☑				3	100%
☑	☑	☑		☑		1	100%
☑	☑	☑			☑	2	100%

Now, let us again assume we have only measured X. Among those exposed to X, a total of 19 have the disease; this leaves 6 who have the disease but are not exposed to X. Given that 50 individuals are exposed to X and 50 are not, our final analysis would be as follows:

$$\mathrm{P(D|exposed\ X)} = \frac{19}{50} = 0.38 = 38\%$$

$$\mathrm{P(D|not\ exposed\ X)} = \frac{6}{50} = 0.12 = 12\%$$

While there is a higher risk of disease given exposure to X, the result is probabilistic rather than deterministic. We observe that individuals who were never exposed to X are still at risk for the disease. These are individuals who were exposed to the sufficient cause that did not include X (that is the sufficient cause that included A, B, and C).

Fully determining who gets the disease and who does not is likely not a productive enterprise. The processes that are involved in each complex disease outcome are likely vast and may involve stochastic processes that are unpredictable. However, identifying those factors that increase the probability of disease at a population level is productive: if we identify one component cause involved in the disease process, we can effectively reduce disease burden associated with that factor and the factors that work in concert with that factor.