Unit 10: Causation

- Criteria for causality
  - » Association vs.
  - » Causation

- Different models

- Different Philosophies

- Hills' Criteria

Dr. A. Sanchez-Anguiano
Epidemiology 6000

Introduction

- Epidemiology: study of the distribution, determinants and deterrents of morbidity and mortality in human populations (Oleckno, 2002)

- Therefore, one of primary goals is to discover "causes".

- Better understanding of "causes" frequently leads to more effective:
  - » prevention and
  - » control measures
- Consequently to a reduction of:
  - » incidence
  - » prevalence or
  - » severity of disease

Introduction (cont.)

- The formulation of etiologic hypotheses most often occurs through the use of descriptive studies. While testing them is the primary function of the analytic study designs.

- Testing an epidemiologic hypothesis involves consideration of the concept of association between a particular exposure and disease.

- **Association** refers to the *statistical dependence* between two variables, the degree to which the rate of disease in those with specific exposure is either higher or lower than the rate of disease among those without the exposure

Introduction (cont.)

- **Definition of HYPOTHESIS**

  A Hypothesis is defined as a
  - » "tentative explanation for
    - – an observation,
    - – phenomenon, or
    - – scientific problem
  - » that can be tested by further investigation."

(Pickett, 2000)

Introduction (cont.)

- An association does not necessarily imply that the observed relationship is one of cause and effect. (rooster crowing at dawn and the rising of the sun; ice cream consumption and drowning during summer. These are non-causal associations).

- Making judgments about causality involves a chain of logic that address *two major areas:* whether the association is *valid* and whether the totality of *evidence* (taken from a number of sources) *supports a judgment of causality*.

Introduction (cont.)

- Assessing *validity* (true relationships between exposure and disease) is a matter of determining the likelihood that *alternative explanations* (chance, bias and confounding) could account for the findings.

- Judging *if the association is causal* extends *beyond validity* of the results of any single study and includes *consideration of other epidemiologic data* as well as the *biologic credibility* of the hypothesis.

Statistical Associations

- Statistically significant associations between exposures and outcomes may be categorized into 3 types.

  - *Spurious.* False. Usually result from sampling error or bias (Random error: alpha=0.05; 5 out of 100 even in well designed studies. Systematic error: bias).

  - *Noncasual.* Real but not causal. Usually represent secondary associations due to confounding factors.

  - *Causal.* Changes in the exposure produce changes in the outcome (in epi we cannot 'prove') ☹ (judgments made using accumulated knowledge).

Causal Associations

1. Association (valid)
   Not due to chance, bias, confounding; evaluated for effect-modification

2. Is the 'valid' association causal? (Is there sufficient evidence to infer that a causal association exists between the exposure and the disease?)

Process of causal inference/judgment of causality requires:
- Valid statistical association
- Assessing whether exposure has actually caused the outcome (evaluation of Hill's criteria).

Chance-uncontrollable force seems to have no assignable cause; unforeseeable & unpredictable process

Evaluating Causal Associations

Causality: A philosophical concept merged with practical guidelines. Can never "prove" causality. Can only infer it.

- The presence of a valid statistical association does not imply that it is a causal one. Among persons 65+,age and sex are associated but one does not cause the other.

- A judgment of causality must be made in the presence of all available data, and reevaluated with each new finding. Never marry a hypothesis. Change your mind as the data change.

- A good scientist has an open mind and maintains objectivity.

Different criteria and philosophical views have been proposed to assess causality.

Disease Etiology. Causation.

- There are several models of disease causation.

- All require the precise interaction of factors and conditions before a disease will occur.

- Models are guidelines that provide a framework for considering causation at a practical level.

- 'Cause' is a concept that is still debated (that is why there are several models to try to explain it.)

'Cause'

- *Cause of disease* is defined as a factor (characteristic, behavior, event, etc.) that *precedes* and influences the occurrence of disease (*not the opposite*), and has a *statistical dependence*. (*Time order, direction & association.* Susser, 1991)
  - » Increase in the factor leads to an increase in disease.
  - » Reduction in the factor leads to a reduction in disease.
 - » There are also inverse relationships.

Models of Causation

- *Models of causation* (examples)
  - » Germ theory: Pasteur, Henle-Kock postulates
  - » Hill's causality criteria
  - » Epidemiologic triad (ecological model).
  - » Multifactorial Model (Rothman's causal pies)
  - » Social-Ecological Model
  - » Holistic Model (W.H.O.)
  - » Wellness Model

Epidemiologic Triad (ecological model)

- The epidemiologic triad (triangle):
  - » Traditional model of infectious disease causation.
  - » It has three components:
    - – *Agent*
    - – *Susceptible host*
    - – *Environment* (brings the other two together; influences the route of transmission of the agent from a source to the host).
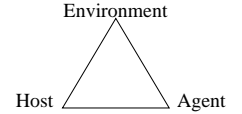

Multifactorial Model / Sufficient – Component Theory (Rothman's causal pies)

- The agent-host-environment model does not work well for some noninfectious diseases.
- A multifactorial model was developed. It is based in the multifactorial nature of causation in many diseases.
- "Component causes" are factors like those intrinsic host factors, the agent and the environmental factors.
- A single component cause is rarely a "sufficient cause" by itself. (ie, *Mycobacterium tuberculosis* is necessary but not sufficient to cause disease).

Multifactorial Model (Rothman's causal pies)

- A particular disease may result from a variety of different "sufficient causes".

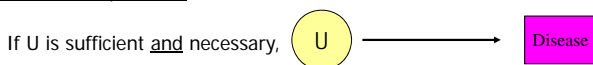Sufficient cause I   Sufficient cause II   Sufficient cause III
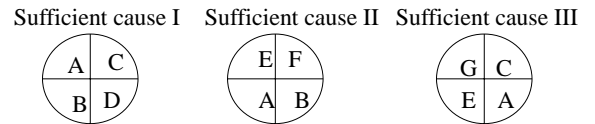
- M. tuberculosis is necessary (A) but not sufficient to cause disease.
- Lung cancer: smoking (B), asbestos (C); both are components, but not necessary causes. (Could occur w/o them)
- To apply this model we do not have to identify every component of a sufficient cause before we can take preventive action. We can block any single component. (ie eliminating smoking (B) would prevent LC in I & II, not in III).
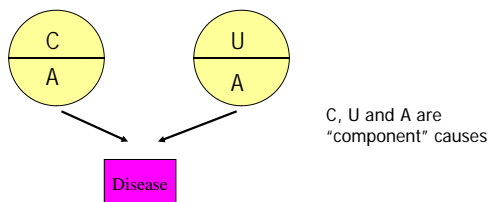
Necessary and sufficient causes: Causal "pies"
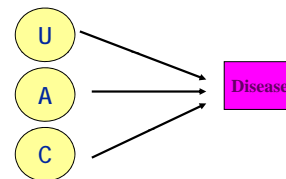
Consider an exposure "U"

If U is sufficient <u>and</u> necessary,
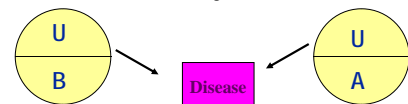
If U is NEITHER sufficient nor necessary,

C, U and A are "component" causes

Causal "pies" (cont.)

If U is sufficient <u>but is not</u> necessary,

If U is not sufficient <u>but is</u> necessary,

Necessary and Sufficient Causes

Disease

	YES	NO	
YES	a	b	SUFFICIENT
NO	c	d	

Exposed
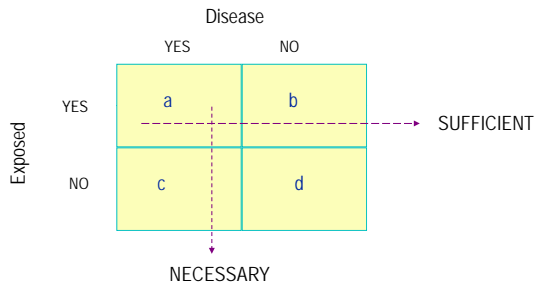
NECESSARY

Necessary:	D does not develop without E --> c = 0
	If D, E=yes
Sufficient:	D always results from E ---> b=0
	If E, D =yes

Examples of Necessary and Sufficient Causes

Sufficient AND Necessary:

Huntington's gene and Huntington's chorea:

No one who has the gene escapes the disease (b=0)
No one with the disease does not have the gene (c=0)

Neither sufficient nor necessary:

Smoking and lung cancer:

Not all smokers get lung cancer (b NE 0)
Not everyone who gets lung cancer smoked (c NE 0)

Examples of Necessary and Sufficient Causes -2

Sufficient but not necessary:

Radiation exposure and thyroid cancer:

Radiation exposure is one (of a number of) way(s) to get thyroid cancer (b=0)
Not everyone with thyroid cancer is exposed to radiation (c NE 0)

Necessary but no sufficient:

Poliomyelitis virus exposure and clinical polio:

Not everyone exposed to the virus developed clinical polio (b NE 0)
Everyone with poliomyelitis was exposed to the virus (c=0)

Different Philosophies of Causal Inference

- Falsification: (Karl Popper - 1959)
  Scientific hypotheses can never be proven or established as true. Therefore, science advances by a process of elimination (falsification)

- Consensus: (Thomas Kuhn - 1962)
  The consensus of the scientific community determines what is considered accepted and what is refuted.

- Inductive-oriented criteria: (Hill 1965)
  Employ a common set of criteria to attempt to distinguish causal from non-causal associations

---

Summary on Causal Theories

1. Multiple philosophies exist for evaluating causality. None are definitive.

2. The set of causal criteria offered by Hill are useful, but are also saddled with reservations and exceptions.

3. Always keep an open mind when evaluating evidence from epidemiologic studies

"I cannot give any scientist of any age better advice than this: the intensity of the conviction that a hypothesis is true has no bearing on whether or not it is true" (Medewar 1979)

Hill's Causal Criteria (guidelines)

- Strength of the association

- Biologic credibility/plausibility

- Consistency with other research

- Temporality

- Dose-response relationship

- Specificity (implies the more diseases an exposure is related to (e.g., smoking), the less likely it is to be causal (faulty)

- Coherence (similar to consistency and plausibility)

- Experimental evidence (not always available or applicable in all settings)

- Analogy (many analogies can be made, few can be generalized to other diseases, e.g., NSAIDs protective for AD, also for ALS? No, risk factor in one abstract presented at the annual American Academy of Neurology meeting, Honolulu, HI (April 2003)

1. Strength of the Association

Pro: The stronger the association, the less likely the relationship is due merely to an unsuspected or uncontrolled confounding variable/bias

Con: Strong but non-causal associations are common

        Example: Non-causal relation between Down's syndrome and birth rank, which is confounded by maternal age

Con: Ratio measures (e.g. RR) may be comparatively small for common exposures and diseases (e.g. smoking and cardiovascular disease), but are causal

Con: When there are many component causes for a disease, each may not have a very strong association with the outcome.

The RR/OR is not always informative in and of itself

Relative risk = 2 means incidence rate of D is twice as high in exposed vs. unexposed.

RR=2   if $p1/p2$ =0.02/0.01 or
                            =0.000002/0.000001 thus the term "relative" risk.

Ex. 1 - incidence rate has increased from 1/100 to 2/100 at risk
(*difference* in risk is 2-1 = 1/100)
Ex. 2 - incidence rate has increased from 1/100,000 to 2/100,000 at risk
(*difference* in risk is 2-1 = 1/100,000)

RR/OR used to measure strength of association and used in judgment of validity and causal nature of an association.

Attributable risk (risk difference, absolute excess) -- difference measure.

---

2. Biologic plausibility of the hypothesis

Pro: A known or postulated biologic mechanism by which the exposure might reasonably alter the risk of developing the disease is intuitively appealing

Con: Plausibility is often based on prior beliefs rather than logic or actual data

Con: What is considered biologically plausible at any given time depends on the current state of knowledge (e.g., tampons and toxic shock syndrome; DES and adenocarcinoma of the vagina):

Kenneth J. Rothman. Epidemiology: An Introduction, Oxford Univ. Press, 2002

How would an epidemiologist apply refutationist thinking to his or her work? If causal mechanisms are stated specifically, an epidemiologist can construct crucial tests of competing hypotheses. For example, when toxic shock syndrome was first studied, there were two competing hypotheses about the origin of the toxin. Under one hypothesis, the toxin responsible for the disease was a chemical in the tampon, so women using tampons were exposed to the toxin directly from the tampon. Under the other hypothesis, the tampon acted as a culture medium for staphylococci that produced the toxin. Both hypotheses explained the relation of toxic shock occurrence to tampon use. The two hypotheses, however, led to opposite predictions about the relation between the frequency of changing tampons and the risk of toxic shock. Under the hypothesis of a chemical intoxication, more frequent changing of the tampon would lead to more exposure to the toxin and possible absorption of a greater overall dose. This hypothesis predicted that women who changed tampons more frequently would have a higher risk of toxic shock syndrome than women who changed tampons infrequently. The culture-medium hypothesis predicts that the women who changed tampons frequently would have a lower risk than those who left the tampon in for longer periods, because a short duration of use for each tampon would prevent the staphylococci from multiplying enough to produce a damaging dose of toxin. Thus, epidemiologic research, which showed that infrequent changing of tampons was associated with greater risk of toxic shock, refuted the chemical theory.

---

3. Consistency of findings with other research

Pro: Due to the "inexact" nature of epidemiologic investigations, evidence of causality is persuasive when several studies conducted by different investigators at different times and in different populations yield similar results

Con: Some effects are produced by their causes only under unusual circumstances

Con: Studies of the same phenomena can be expected to yield different results simply because they differ in their methods and from random errors.

---

4. Temporality/Temporal Sequence

Pro: By definition, a cause of disease must precede onset of the disease. <u>An absolute must</u> (the only one).

Not really a con, just a problem:

    The existence of an appropriate time sequence can be difficult to establish (e.g. does stress lead to overeating or does overeating lead to stress?

5. Dose-response Relationship/Biologic Gradient

Pro: Logically, most harmful exposures could be expected to increase the risk of disease in a gradient fashion (e.g. if a little is bad, a lot should be worse)

Note: Some associations show a single jump (threshold) rather than a monotonic trend

Note: Some associations show a "U" or "J" shaped trend (e.g. alcohol consumption and mortality; maternal age at time of birth)