



MAKING CAUSAL INFERENCES

ORGANIZED BY THE LAWYERS AT ANAPOL SCHWARTZ. © 2009 ALL RIGHTS RESERVED.

CONTACT LAWYER: BARRY HILL, ESQUIRE
CALL: (215) 735-0364
EMAIL: BHILL@ANAPOLSCHWARTZ.COM
READ MORE INFORMATION ONLINE AT: www.anapolschwartz.com

DISCLAIMER: This information is not intended to replace the advice of a doctor. Please use this information to help in your conversation with your doctor. This is general background information and should not be followed as medical advice. Please consult your doctor regarding all medical questions and for all medical treatment.



REVIEW OF SOME EPIDEMIOLOGICAL AND STATISTICAL CONCEPTS

- **Epidemiological concepts of**
 - disease incidence and prevalence
 - relative risk
- **Statistical concepts of**
 - p-values
 - confidence intervals

Incidence and Prevalence

- **Incidence Rate**
 - number of new disease cases per person-year of the population of interest
e.g., 13 myocardial infarctions per 1,000 male smokers per year
 - the incidence rate represents the “risk” of developing a disease
- **Prevalence**
 - number of existing disease cases per person in the population of interest at a specific instant in time
e.g., 45 influenza cases per 1,000 school students on September 1st
 - a proportion, not a rate

Relative Risk (RR)

- **ratio of the risk (i.e., incidence rate) in an exposed population to the risk in an otherwise similar unexposed population**
 - e.g., if the incidence of heart attacks in male smokers is 13 per 1000 person-years, and in non-smoking men is 5 per 1000 person-years, the RR of heart attack associated with smoking is 2.6
- **note that a RR of 1 means no association**
- **also known as “risk ratio”**

RELATIVE RISK (RR)

- **Relative Risks can be estimated using various epidemiological study designs, including**
 - clinical trials
 - cohort studies
 - case-control studies
(here “odds ratios” (OR) are used as estimates of RR)
- **the RR is the measure of effect, the strength of an association**
- **it is an indicator of the biological significance of an association between an exposure and a disease**

Statistics: p-values

- **statistics are used to indicate the likelihood that a study’s findings are due to chance**
- **in data analysis, we can calculate not only the RR, but also the probability that the association observed was due to chance, e.g.,**
 - RR = 2.6, $p = 0.001$
i.e., 0.1% probability that the observed RR was a chance finding
- **p-values are affected primarily by 3 factors**
 - the magnitude of the difference in risk between the exposed and unexposed groups
 - variability within the two groups
 - the number of study subjects

Statistics: confidence intervals

- **a bridge between the measure of effect and p-value**
- **indicates both the strength of association and the likelihood that the difference in risk between the groups is not coincidental**
- **95% confidence interval, e.g.,**

RR = 2.6 95% CI: 1.8 - 3.7 $p = 0.001$

Thus, there is 95% probability that the true population RR lies between 1.8 and 3.7.

RELATIVE RISK VS. P - VALUE

- **measure of effect (RR) ~ biological significance**
 - estimates ratio of disease risks between groups
 - if RR indicates a strong association
 - more likely that exposure truly associated with disease
 - less likely that difference is due to chance alone
- **p-value ~ statistical significance**
 - does not indicate the biological significance of the exposure
 - measures role that chance may have played

INTRODUCTION TO CAUSAL INFERENCE

- **Why be concerned with cause?**
 - so that we can intervene and prevent disease
- **basic definition of cause**
 - exposure that leads to new cases of disease
 - remove exposure and some cases do not occur (disease rate drops)
- **do not need to understand all causal factors to prevent disease**



ASSOCIATION VS. CAUSATION

- **Association is simply an identifiable relationship between an exposure and disease**
 - e.g., Coronavirus is isolated more frequently from individuals with diarrhea than those without
- **Implies that exposure might cause disease**
- **Exposures associated with a difference in disease risk are often called “risk factors”**

Association vs. Causation

- **Causation implies that there is a true mechanism that leads from exposure to disease**
 - e.g., long-term heavy smoking causes myocardial infarction
- **Finding an association does not make it causal**
 - e.g., hospital stays are associated with an increased mortality rate, but this does not mean they cause death

BRADFORD HILL’S CRITERIA FOR CAUSAL INFERENCE

- **the list:**
 - Consistency of findings
 - Strength of association
 - Biological gradient (exposure-effect)
 - Temporal sequence
 - Biological plausibility
 - Coherence with established facts
 - Specificity of association

A MORE COMPLETE DESCRIPTION OF EACH ONE . . .

Consistency of Effect

- Relationships that are demonstrated in multiple studies are more likely to be causal, i.e., consistent results are found
 - in different populations,
 - in different circumstances, and
 - with different study designs.

Strength of Association

- Strong associations are less likely to be caused by chance or by bias
 - a strong association is one in which the relative risk is
 - very high ($\gg 1$), or
 - very low ($\ll 1$)

Biological Gradient (exposure-response)

- Changes in exposure are related to a trend in relative risk

Mortality Rates

Dose (cigs/day)	Study 1	Study 2	Study 3
<10	1.3	1.8	1.4
10-20	2.8	2.3	2.4
>40	4.7	3.7	6.3

Temporal Sequence

- Exposure must precede disease
- In diseases with latency periods, exposures must precede the latent period
- In chronic diseases, often need long-term exposure for disease induction

Biological Plausibility & Coherence with Established Facts

- **The proposed causal mechanism should be biologically (etiologically) plausible**
- **Causal mechanism proposed must not contradict what is known about the natural history and biology of the disease, but**
 - the causal relationship may be indirect
 - data may not be available to directly support the proposed mechanism
 - must be prepared to reinterpret existing understanding of disease processes in the face of new findings

Specificity of the Association

- **A cause leads to a single effect**
- **Easier to support causation when associations are specific**
- **But, obviously not always true**
 - many exposures cause multiple diseases,
e.g., smoking causes many diseases including heart disease, lung and other cancers, emphysema
 - so some call this criterion “useless and misleading”

Causal Inference

- **No single study is sufficient for causal inference**
- **Causal inference is not a simple process**
 - consider weight of evidence, using Bradford Hill’s criteria
 - always requires judgement and interpretation, no cookbook method
 - some consider causal inference to be in the public policy domain, rather than the scientific domain
- **No way to prove causal associations**

SOME NOTES FROM BRADFORD HILL

A caveat:

“None of my . . . [criteria] can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a sine qua non”

But also a need to act without “proof”:

“All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time.”

