Overview of analytic study designs

Descriptive studies

Lack hypotheses specified in advance

Often used to discover associations, whether causal or not

Types

1. Case Report -Surveillance of rare events
    - A case report is a detailed report of the symptoms, signs, diagnosis, treatment, and follow-up of an individual patient. Case reports may contain a demographic profile of the patient, but usually describe an unusual or novel occurrence.
2. Case Series -Description of unusual case group
    - Case reports consist of collections of case reports on the treatment of several patients. Because this is a report of cases and use no control groups with which to compare outcomes, they have no statistical validity.
3. Ecological Studies – typically Correlation studies
    - In a typical ecological study, two ecological variables are contrasted to examine their possible association.
    - Typically this is an ecological measure of exposure and an aggregate measure of disease or mortality are compared

Analytic studies

A. General characteristics
    1. Have hypotheses specified in advance
    2. Usually intended to establish causal associations
    3. Main types of analytic studies often result in data which can, in their simplest form, be organized as follows:

		Outcome: Bad	Outcome: Good
Exposed to factor of interest?	Yes	a	b
	No	c	d

where a, b, c, and d are number of subjects in each cell

    4. Further classification of the **study design** depends on:
        a. The basis on which the main comparison groups were formed at the beginning of the study
        b. Whether the study involved a Period of follow-up between exposure and outcome

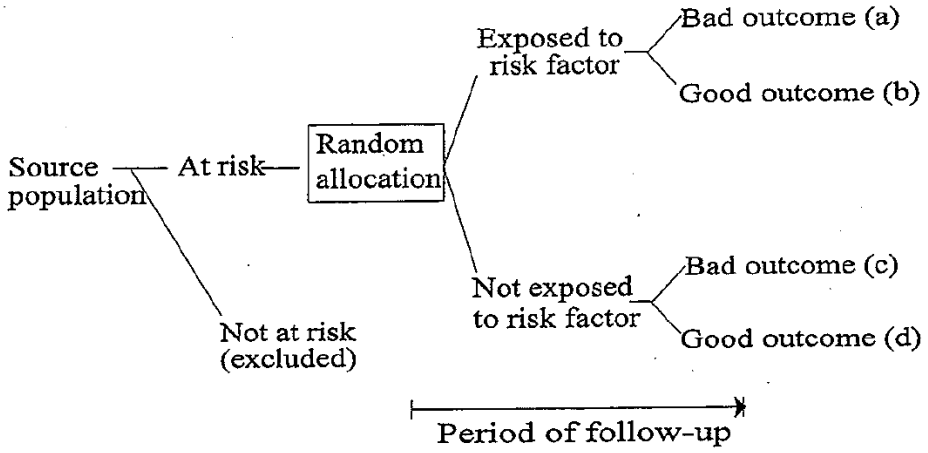
B. True experimental (randomized) studies

Generally have both of the following characteristics:

1. Investigator has control over which study subjects are exposed to the factor of interest, and makes exposure assignments to serve the purposes of research

2. Two or more comparison groups are formed by <u>random assignment</u> of subjects to groups which differ on degree of exposure to a factor of interest: e.g., a vaccine

3. Prototype design for a **<u>Randomized Trial</u>** (experiment):



4. Generally considered the strongest design for establishing causal relationships

5. Characteristics mentioned in (1) and (2) above do not always go together, leading to some differences of opinion regarding proper use of the term "experiment"

   a. Intervention studies without random assignment are sometimes termed "quasi-experiments." We shall treat them as non-experimental studies.

   b. Studies in which the investigators had no control over determining exposure status of study subjects, but which nonetheless involved random (or effectively random) assignment of subject to exposure groups are sometimes termed "natural experiments." We shall treat them as experiments if the allocation was truly at random and as non-experimental studies otherwise.

C. Non-experimental studies: no random assignment

1. **Before-after** (single group only)

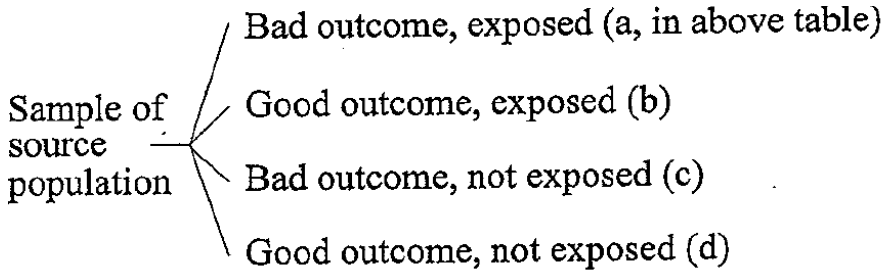
   a. Involves comparing status of subjects at baseline and following some known exposure or event

   b. Strengths: inexpensive

   c. Weaknesses: often unclear what would have happened in the absence of exposure

2. **Cross-sectional**

   Prototype:



   a. Classification of the status of study subjects on exposure and outcome refers to the same point in time

   b. Often use "naturalistic" sampling--i.e., sample not chosen to achieve any particular mix of exposed vs. non-exposed subjects or good vs. bad outcomes, but to represent a "cross-section" of a population of interest

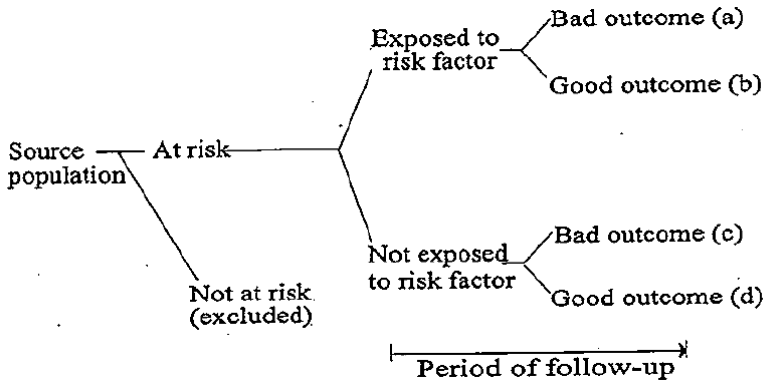
   c. Strengths: when subjects selected as described in (b), can yield data on prevalence of outcomes and exposures

   d. Weaknesses

      - May be unclear whether exposure preceded outcome
      - Associations are between exposure and prevalence, not incidence
      - Naturalistic sampling is statistically inefficient: can show that for any given number of subjects, one of the other non-experimental designs (cohort or case-control--see below) will always have at least as much statistical power (and usually more power) to find a hypothesized effect if it exists

3. **Cohort (follow-up) study**

Prototype:



a. Main comparison groups are formed on the basis of exposure status at the beginning of a period of follow-up

b. Relative sizes of exposed and non-exposed groups may not necessarily reflect exposure frequency in the source population: e.g., may include all exposed persons and only a sample of those not exposed

c. Resembles an experiment, except that there is no random assignment of subjects to exposure groups; in fact, investigator usually has no control at all over which subjects are exposed and which are not

d. **Prospective cohort study**: period of follow-up occurs after (in calendar time) ascertainment of exposure status for each subject, as investigators watch and wait for outcome events to occur in both groups

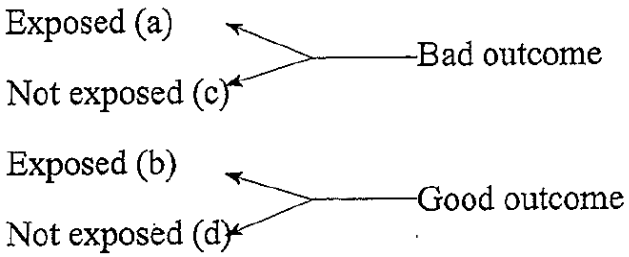
e. **Retrospective cohort study**: period of follow-up has already occurred before study is initiated
   - Essentially a reconstruction of a cohort study which has already taken place
   - Generally much cheaper and faster than a prospective cohort study
   - Depends on having accurate exposure data on subjects at some time in the past, and accurate follow-up on nearly all subjects

f. Strengths
   - Time sequence of exposure and outcome unambiguous
   - Can directly measure the risk of a bad outcome
   - Can study multiple outcomes of a single exposure in the same study

g. Weaknesses
   - Inefficient for rare outcomes
   - Prospective cohort studies can be expensive if they require large samples or long follow-up periods

4. **Case-control study**

a. Proceeds backward from effect to cause:



b. At start of study, main comparison groups are formed on the basis of outcome status

c. Relative number of subjects with good vs. bad outcomes usually does not reflect the real frequency of these events in the source population; often, all available subjects with bad outcomes (cases) are included and only a small fraction of the non-cases

d. Under certain assumptions, the odds ratio (= ad/bc) from a case-control study provides an excellent estimate of relative risk

e. Strengths:

- Efficient for rare outcomes
- Permits study of multiple possible causes of an outcome of interest
- Usually relatively inexpensive and fast

f. Weaknesses

- Can't directly obtain the absolute risk (i.e., incidence) of a bad outcome
- Selection of appropriate controls may be tricky
- Self-reported exposure data may be subject to biased recall due to subjects' knowledge of their outcome status

References

1. Rothman KJ & Greenland S. Modern epidemiology. Lippincott- Raven, 1998 (Chapter 5).
2. Kleinbaum DG, Kupper LL, and Morgenstern H. Epidemiologic research: Principles and quantitative methods. Belmont, CA: Lifetime Learning Publications, 1982, Chapters 4 and 5.
3. Bailar JC III, Louis TA, Lavori PW, Polansky M. A classification for biomedical research reports. N Engl J Med 1984; 311:1482-1487.
4. Fleiss JL. Statistical methods for rates and proportions (2nd edition). New York: Wiley and Sons, 1981 (Pages 20-24 and Chapters 5-7).
5. Kramer MS, Boivin J-F. Toward an "unconfounded" classification of epidemiologic study design. J Chron Dis 1987; 40:683-688.
6. Miettinen OS. Striving to deconfound the fundamentals of epidemiologic study design. J Clin Epidemiol 1988; 41 :709-713.
7. Greenland S, Morgenstern H. Classification schemes for epidemiologic research designs. J Clin Epidemiol 41:715-716.
8. Feinstein AR. Directionality in epidemiologic research. J Clin Epidemiol 41 :705-707.
9. Kramer MS, Boivin J-F. The importance of directionality in epidemiologic research design. J Clin Epidemiol 1988; 41:717-718.
10. Feinstein AR. Classification of epidemiologic research. J Clin Epidemiol 1988; 41:805.
11. Kramer MS, Boivin J-F. Response. J Clin Epidemiol 1988; 41:806.

SHORT GLOSSARY OF STUDY DESIGN TERMS

Design name from outline	Approximate synonyms
Descriptive study	Hypothesis-generating study
Analytic study	Hypothesis-testing study
Non-experimental study	Observational study
Before--after study	Pretest--post-test study
Cross-sectional study	Correlational study Prevalence study Survey
Cohort study	Follow-up study Longitudinal study Incidence study Prospective study
Prospective	Concurrent cohort study Futuristic cohort study
Retrospective	Historical cohort study Retrospective prospective study
Case-control study	Case-referent study Case-compeer study Causal-comparative study Retrospective study
Experiment	Randomized controlled trial Clinical trial

COMPARATIVE STRENGTHS OF COHORT AND CASE-CONTROL STUDIES

	Case-control	**Cohort**
Strengths:	• Well suited to rare outcomes	• Well suited to rare exposures
	• Can easily study multiple exposures with same comparison groups	• Can easily study multiple outcomes with same comparison groups
	• Efficient if long delay between exposure and outcome	• Provides information about absolute risk of outcome among both exposed and unexposed persons
	• Often less expensive and time-consuming than other competing designs	
Weaknesses:	• Inefficient for rare exposures	• Inefficient for rare outcomes
	• Not well suited to study of multiple outcomes	• Not well suited to study of multiple exposures
	• Unable to provide data on absolute risk	• Difficult if long delay between exposure and outcome*
	• Time sequence of exposure and outcome can be unclear	• Can be costly*

*Less a problem in retrospective cohort studies than in prospective cohort studies.