

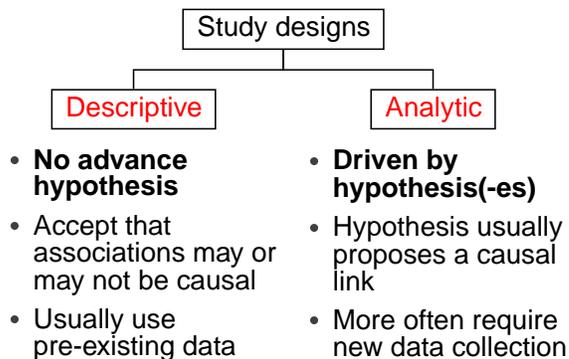
Overview of Study Designs

Seattle Epidemiology and
Biostatistics Summer Session
June, 2004

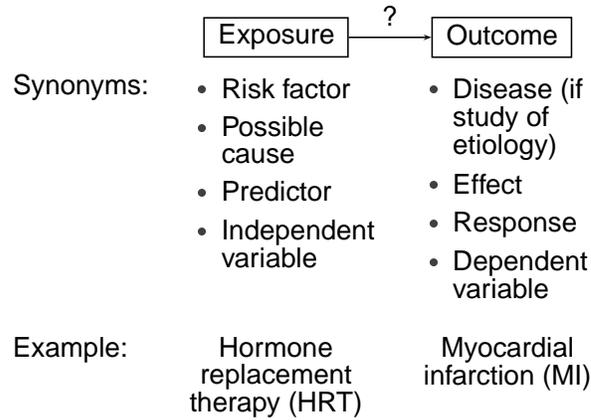
Introduction

- **Study design:** plan for selection of study subjects and collection of data on them
- Possibilities are infinite
- Focus here on:
 - Major distinctions among designs
 - Designs most commonly used
 - Generic strengths and weaknesses of key designs
- For present purposes, assume potential study subjects already screened for eligibility and willingness to participate

Descriptive vs. analytic studies



Typical hypothesis in an analytic epidemiologic study

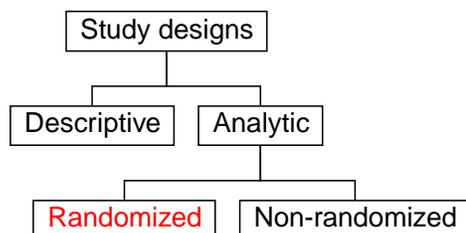


Data layout common to many analytic studies

| | Outcome | |
|----------|----------|----------|
| Exposed? | Bad | Good |
| Yes | <i>a</i> | <i>b</i> |
| No | <i>c</i> | <i>d</i> |

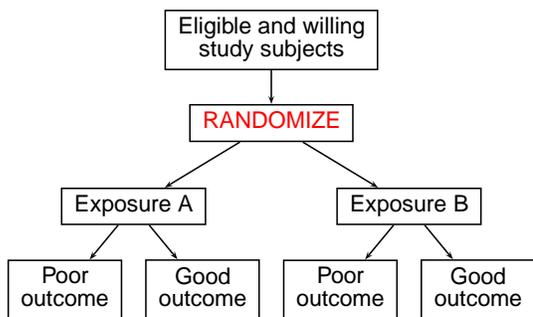
where a, b, c, d are number of study subjects in each cell

Types of analytic studies

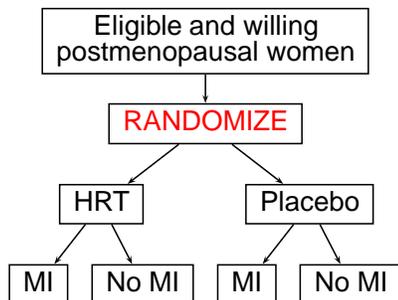


Randomized controlled trial (RCT)

Two-arm randomized trial



Randomized trial: HRT and MI



Key features of a randomized trial

- An *intervention* study: investigator dictates which study subjects are exposed
- **Random assignment** used to decide on each subject's exposure
- A *prospective* study: all outcomes of interest occur after study has begun
- Many design variations

Strengths and weaknesses of RCTs

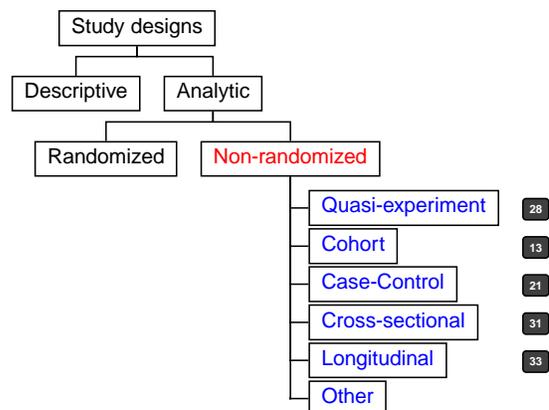
Strengths

- Superior control over *confounding factors*, even if unknown or hard to measure
- Exposure clearly precedes outcome
- Can estimate incidence in both groups
- Easy to study several outcomes

Weaknesses

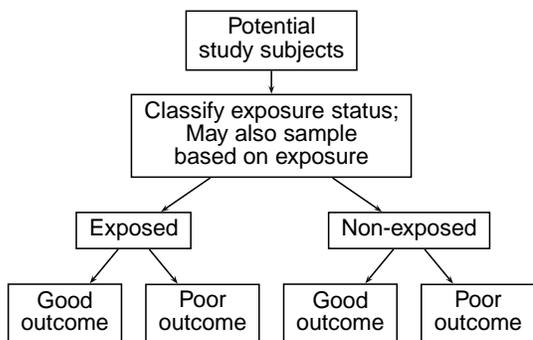
- Not always possible or ethical to manipulate exposure at random
- Inefficient for rare or delayed outcomes

Non-randomized analytic studies

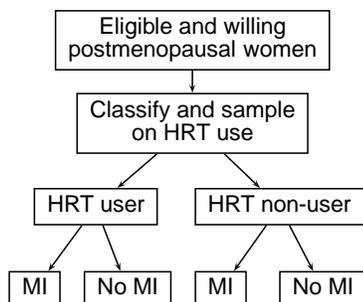


Cohort study

Generic cohort study



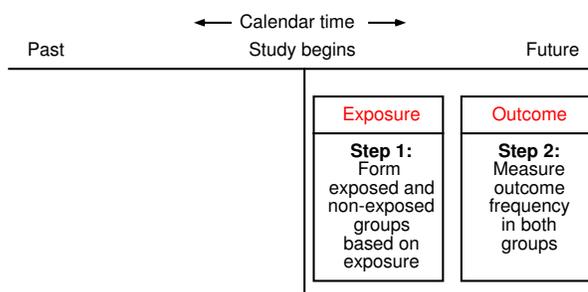
Cohort study of HRT and MI



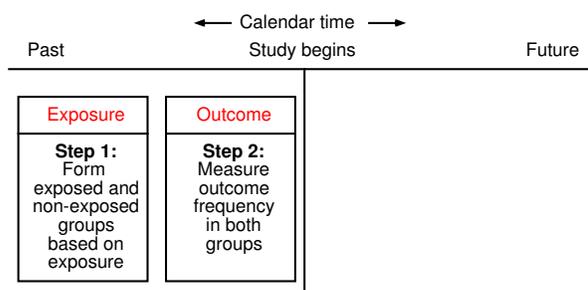
Key features of a cohort study

- An *observational* study: investigator observes, but does not control, which study subjects are exposed
- Relative sizes of exposed and non-exposed groups need not necessarily reflect frequency of exposure in source population
- Incidence of outcomes monitored over time in exposed and non-exposed groups
- Two subtypes:
 - Prospective
 - Retrospective

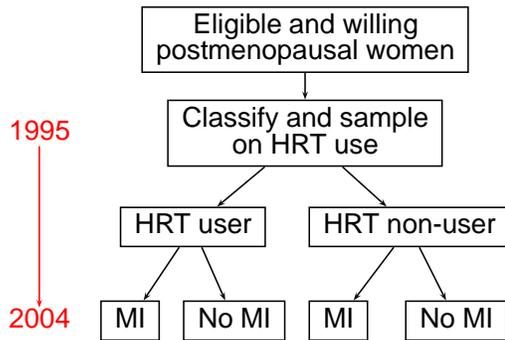
Prospective cohort study



Retrospective cohort study



Retrospective cohort study, begun in 2004

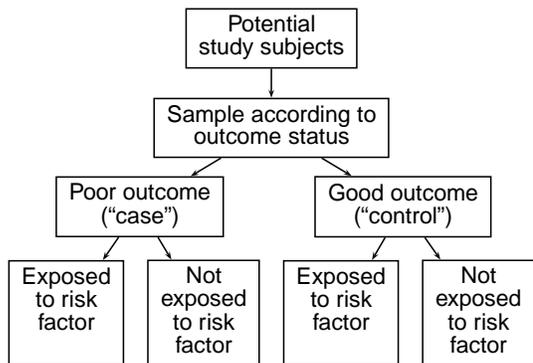


Strengths and weaknesses of cohort studies

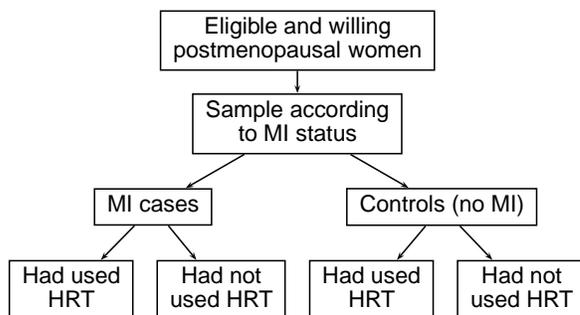
| Strengths | Weaknesses |
|---|--|
| <ul style="list-style-type: none"> • Exposure known to precede outcome • Can estimate incidence in both groups • Easy to study 2+ different outcomes • Efficient for rare exposures | <ul style="list-style-type: none"> • Inefficient for rare outcomes • If prospective, can be costly for large samples or delayed outcomes |

Case-control study

Generic case-control study



Case-control study of HRT and MI



Key features of a case-control study

- Relative sizes of case and control groups usually do not reflect frequency of outcome in the source population. Typically:
 - All available cases included
 - Only a sample of many non-cases included
- Hence cannot directly estimate incidence from case-control data
- Nonetheless, can nearly always estimate *relative* incidence:

$$RR \approx Odds\ Ratio = OR = ad/bc$$
 (See text for details)

Strengths and weaknesses of case-control studies

| Strengths | Weaknesses |
|--|---|
| <ul style="list-style-type: none"> • Efficient for rare or delayed outcomes • Usually relative quick and inexpensive • Easy to study 2+ different exposures | <ul style="list-style-type: none"> • Cannot estimate incidence directly • Choice of appropriate controls can be difficult • Recall bias possible if exposure self-reported |

Cohort sampling

Full population (unobserved):

| | MI | No MI | |
|--------|-----|-------|--------|
| HRT | 60 | 1,940 | 2,000 |
| No HRT | 400 | 7,600 | 8,000 |
| | 460 | 9,540 | 10,000 |

$$RR = \frac{60/2,000}{400/8,000} = 0.6$$

Cohort sampling:

| | MI | No MI | |
|--------|----|-------|-------|
| HRT | 15 | 485 | 500 |
| No HRT | 25 | 475 | 500 |
| | 40 | 960 | 1,000 |

$$RR = \frac{15/500}{25/500} = 0.6$$

Case-control sampling

Full population (unobserved):

| | MI | No MI | |
|--------|-----|-------|--------|
| HRT | 60 | 1,940 | 2,000 |
| No HRT | 400 | 7,600 | 8,000 |
| | 460 | 9,540 | 10,000 |

Case-control sampling:

| | MI | No MI | |
|--------|-----|-------|-----|
| HRT | 60 | 94 | 154 |
| No HRT | 400 | 366 | 766 |
| | 460 | 460 | 920 |

$$RR \approx OR = \frac{60 \times 366}{94 \times 400} \approx 0.58$$

Quasi-experiment

Quasi-experiment

- A *non-randomized* intervention trial
- Investigator can manipulate exposure. . .
- . . . but does not use randomization to assign subjects to different exposure groups
- More vulnerable to confounding than an RCT

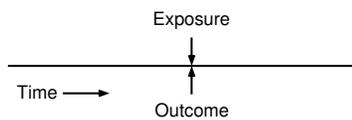
RCT vs. quasi-experiment vs. cohort study

| Design | Formation of exposure groups |
|------------------|--|
| RCT | Investigator uses randomization to dictate who will be exposed |
| Quasi-experiment | Investigator dictates who will be exposed, but does not use randomization to do so |
| Cohort study | Investigator merely a passive observer: classifies (but does not manipulate) exposure status |

Cross-sectional study

Cross-sectional study

- Exposure and outcome assessed as of the same time or time period for each study subject:

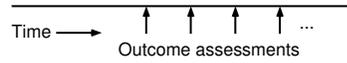


- Hence cannot tell from study data alone whether exposure preceded outcome

Longitudinal study

Longitudinal study

- Involves multiple measurements of outcome (and possibly exposure) over time for each subject:

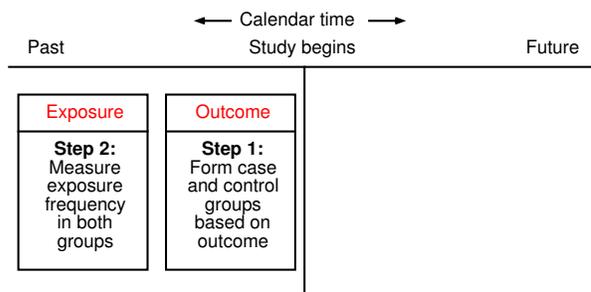


- May allow within-person comparison of outcomes after exposed and non-exposed periods

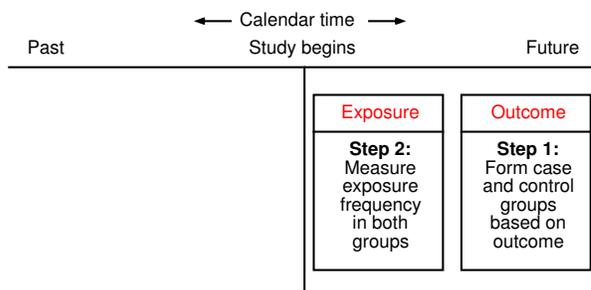
Example: Effects of different diets on body weight

12

Retrospective case-control study



Prospective case-control study



Introduction to Epidemiologic Methods — Summer, 2004

Discussion Questions: Study Designs

1. Not long ago, the Food and Drug Administration was considering whether to ban phenylpropanolamine (PPA), an ingredient in many widely used non-prescription cold remedies. Concerns have been raised that PPA use may increase the risk of hemorrhagic stroke, a rare but serious neurologic event. What general kind of study design would you consider most suitable for testing whether this concern is well-founded? Briefly explain your choice.
2. Advances in in-vitro fertilization (IVF, or “test-tube” conception) have led to growing use of the technique worldwide. In Sweden, about 2% of babies are born following IVF. A group of Swedish investigators wished to determine the extent to which IVF is associated with several adverse postnatal and early childhood outcomes, particularly low birth weight, cerebral palsy, congenital malformation, mental retardation, and chromosomal abnormality. Data sources available to them included:
 - The Swedish Medical Birth Registry, which includes birth certificate data on all children born in Sweden
 - Records of the National Board of Health and Welfare, which receives reports from all 14 IVF clinics in Sweden about which mothers underwent IVF and when they had the procedure
 - Reasonably accurate diagnosis data on patients treated at all 26 childhood disability centers in Sweden, which provide care to children with the various congenital and developmental abnormalities of interest under a national health care system

All of these data sources have been maintained since 1982 or earlier and are expected to continue. Sweden also assigns a unique personal identification number to each citizen, including mothers and newborns, that can be used to link data across data sources.

If you were a member of the investigative team, what basic epidemiologic study design would you suggest as best-suited to the research aims and data resources available? Briefly justify your choice.

3. A recent study sought to identify factors associated with procedural errors during surgery that involve leaving a surgical instrument or sponge inside the patient. Records of a large malpractice insurer were reviewed to identify all such claims during a 6-year period in a certain state. When a qualifying claim was found, administrative records were reviewed at the hospital where the ill-fated surgery had been performed. Ten other patients who had undergone the same surgical procedure at the same hospital within the preceding 6 months were identified, omitting any who also appeared on the claims list. This list of ten patients was then sorted into random order, and the first four patients on the randomized list for

whom complete records could be found were included in the study for comparison with patients on the claims list.

- (a) How would you classify the basic study design?
- (b) Which of the following quantities would you expect the investigators to be able to estimate from this study?
 - i. The incidence of a retained instrument or surgical sponge among patients who underwent emergency surgery
 - ii. Ratio of the incidence of a retained instrument or surgical sponge among patients who underwent *emergency* surgery to the incidence of a retained instrument or surgical sponge among patients who underwent *non-emergency* surgery, assuming they collected information about whether each operation was performed on an emergency basis