



Research Design and Development

Part II: Observational Study Designs

Christine A Jesser, ScD

Senior Epidemiologist

Seton Family of Hospitals

Office of Research Administration

Observational Study Designs



Three basic types, classified according to the sampling mechanism:

- Cohort

- Cross Sectional

- Case-control

{
– Prospective
– Retrospective



Special situation – case report or case series

- Examination of a single case or a small number of cases to describe clinical course of disease/condition
- Can serve to generate hypotheses
- There is no federal regulation or guidance that determines how many cases constitute “research”
- Institution responsible for defining number of cases examined that constitute “research”
- ORA finalizing policy for Seton researchers

Observational Studies - Review

- When should observational studies be performed?
 - Cultural issues make trials difficult in certain areas, including Surgery or Anesthesiology (and blindness is often impossible)
 - Exploratory (generate hypothesis) or preliminary study
 - When exposure of interest cannot be manipulated by the investigator
 - When testing orphan drugs or interventions for very rare diseases (FDA can accept) or Phase IV trials
 - Studies about accuracy of diagnostic tests
 - Outcomes research

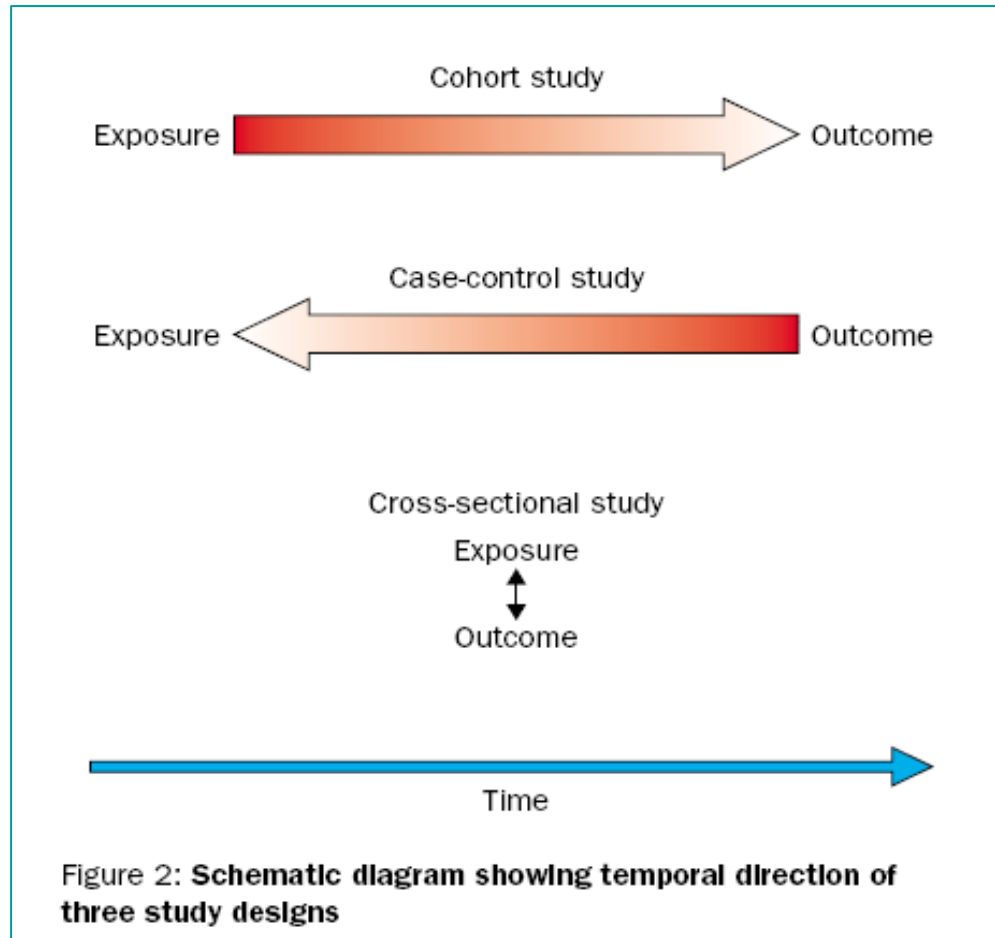
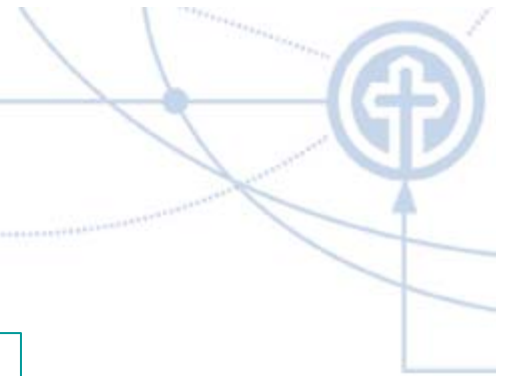
Study Timing



Retrospective vs. Prospective:

- ***Timing of measurement of exposure and covariates***
(retrospective = measurement obtained before outcome occurred; prospective = measurement obtained after outcome occurred)
- ***Timing of identification of subjects with regards to end of follow-up***

Epidemiologic Study Designs



Cross-Sectional Study



The observation of a population, or a representative subset of this population, at a defined point in time

Example

Study Aim:

- A) Estimate the frequency of RSV infection in children admitted with complains of upper respiratory infection at a large tertiary pediatric hospital ED;
- B) Assess the association between RSV (as compared to Influenza) and age groups and temperature

Cross-Sectional Study

Measures Disease Frequency



- **Prevalence:** proportion of subjects who acquired the disease at some point in the past and remain sick and alive in the present
- Prevalence is a function of the duration, fatality, and incidence of disease
- Notice any proportion can be estimated in the 2X2 table

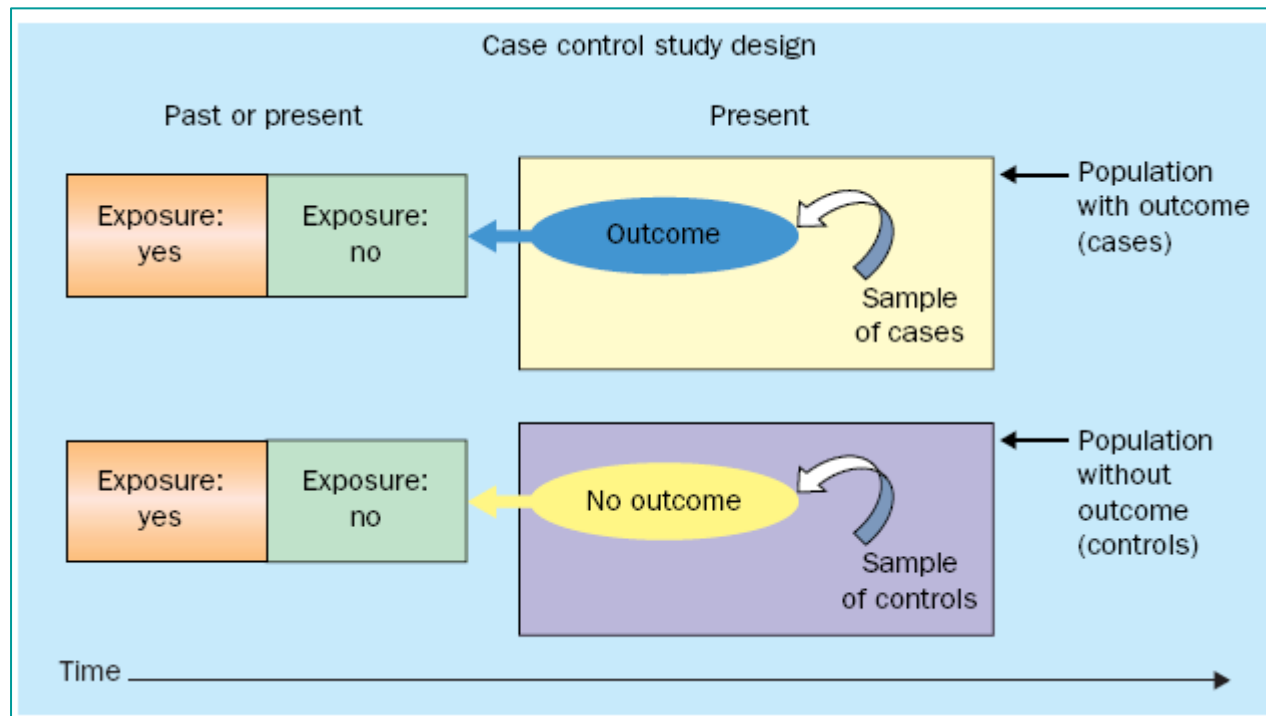
	Exp	Not Exp	
Disease	a	b	nD
No Disease	c	d	
	E	nE	N

Cross-Sectional Study Strengths and Weaknesses



- Measures association with duration and occurrence (not only occurrence).
- No time direction between exposure and outcome (which came first: chicken or the egg?)
- Impractical for rare diseases (requires too large a sample size)

Case-Control Studies



Schulz & Grimes, 2002 ([www](#)) ([PDF](#))

Case-Control Study



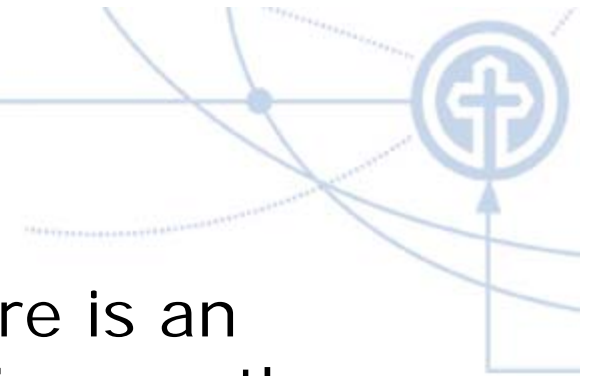
First, Select:

*Then, measure
past Exposure*

	Cases (with Dz)	Controls (without Dz)
Were Exposed	a	b
Were NOT Exposed	c	d
Total	a + c	b + d
Proportions Exposed	$\frac{a}{a + c}$	$\frac{b}{b + d}$

Case-Control Study

- Thus, in a case-control study, if there is an association of an exposure with a disease, the prevalence of history of exposure should be higher in persons who have the disease (cases) than in those who do not (controls).



Case-Control Study



	CHD	Controls
Smoke	112	176
Do not smoke	88	224
Total	200	400
% Smoking Cigarettes	56.0	44.0

Notes on Case-Control Studies

- IF we use only data from a case-control study, we cannot estimate the prevalence of the disease.
 - The number of controls selected is up to the investigator
 - The optimal ratio depends on a couple factors, but is usually 1 or 2 controls per case

Selection of Cases

****need to specify selection criteria****

- Cases should be “representative”, i.e. registries vs. hospitals
 - Cases selected from hospitals may track with other features associated with admission to that hospital
 - Disease registries are valuable for non-biased case selection.

Selection of Cases



- Incident vs. Prevalent Cases
 - In studies of etiology, preferable to use incident cases.
 - Why?
 - Factors associated with prevalent disease may be associated with survival rather than to the development (incidence) of the disease.
 - E.g. cases may die soon after disease onset and thus, not be represented in a study of prevalent cases.

Definition of Controls

- Controls should be subjects who would be included as cases in our study if they would have developed the outcome.
- In other words, they constitute the Study Base (a population from which the cases eligible for the study were generated)



Selection of Controls

- When cases are selected from health service (or registries, or insurance) with catchment area circumscribed:
 - Random sample from same population can be obtained for instance via random digit dialing.





Selection of Controls

- When cases are selected from health services (or registries, or insurance) with catchment area not- circumscribed:
- Friends or neighbors (as long as exposure, like smoking, cannot contribute to establish friendship),
- Relatives (siblings of the cases), can be problematic for some exposures like parent's characteristics or some genetic factors.
- Visitors of patients
- Hospital controls

Selection of Controls



- Controls should have (take home message):
 - Same selective factors of the cases.
Remember! Catchment area of hospitals is different for different diseases due to specialists, etc.
 - Same information quality of cases (quality of information of community control < healthy visitor < sick patient)
 - It is safer to select controls with more than one disease (or two disease groups)

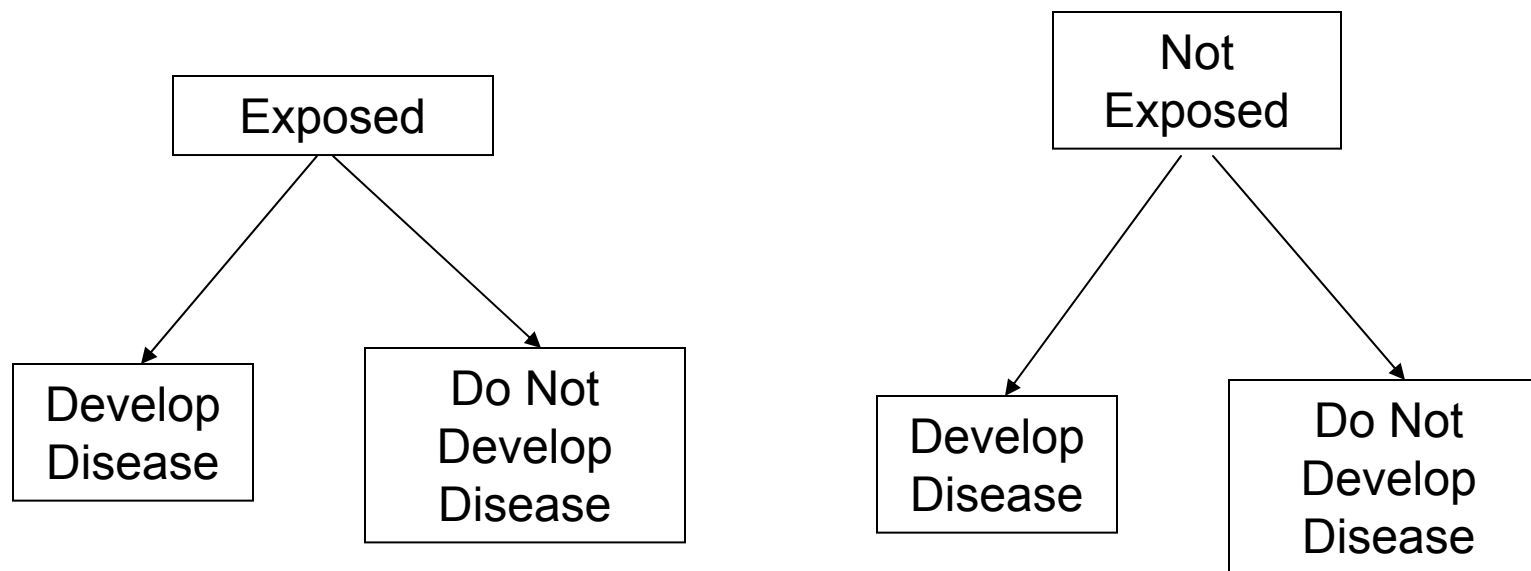
Case-control Study

Strengths and Weaknesses



- Highly susceptible to bias:
 - Difficult to select good controls (prone to confounding and selection bias)
 - Often measures exposure after the outcome occurred (prone to information bias)
 - Cannot directly estimate true causal parameter of interest (“incidence of outcome in exposed”)
- Can be very biased in fatal diseases
- Less expensive and good to study rare disease or diseases with a long induction time
- Good for “preliminary” studies

Design of a Cohort Study



Cohort studies START with an observation period or an exposure.

Cohort Study



Then follow to see if:

	Disease Develops	Disease Does Not Develop	Totals	<i>Incidence Rates</i> of Disease
<u>First, Select</u>				
Exposed	a	b	a + b	$\frac{a}{a + b}$
NOT Exposed	c	d	c + d	$\frac{c}{c + d}$

NOTE: Cohort studies allow for precise measurement of exposure.

Cohort Study (smoking and CHD)



Then follow to see if:

	Disease Develops	Disease Does Not Develop	Totals	<i>Incidence per 1,000 per Year</i>
<i>First, Select</i>				
Smoke	84	2,916	3,000	28.0
Do Not smoke	87	4,913	5,000	17.4

Because these are NEW cases, we can establish a temporal relationship between the exposure (smoking) and CHD.

Measures Disease Frequency Cohort Study



- Risk or proportions (risk ratio; difference of proportions)
- Rate or Hazard (hazard ratio)
- Length of time or duration (difference or relative change)
- Means (difference of means)

Cohort Study

Strengths and Weaknesses



- Establishes temporal association
- Measures exposure before the outcome occurred (less prone to information bias)
- May be expensive and unfeasible, particularly when studying rare diseases

Next time...

- Please join the Office of Research each month for lectures on research design, development, medical ethics, contracting, grants management, IRB processes and procedures, and more.

Thank you!

