



UNIVERSITY OF
SASKATCHEWAN

Study Design & Measures of Association

Dr. Daniel Fuller
Department of Community Health & Epidemiology

Definition of bias

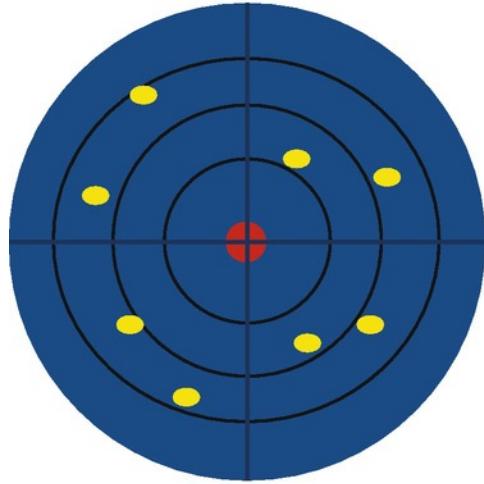
Bias is a systematic error introduced by the study design, analysis, and interpretation

Error is a deviation of your findings from the truth

Random error (not bias) is caused by the fluctuation of sampling (your cohort members are different, every time you sample from your source population)

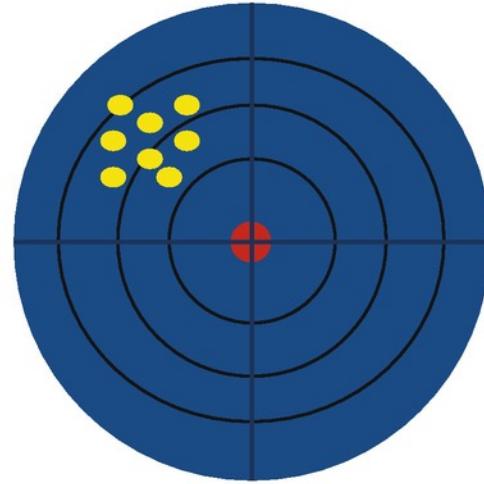
Random vs. systematic error

If we repeated studies many times and make estimates (e.g. mean, proportion)...



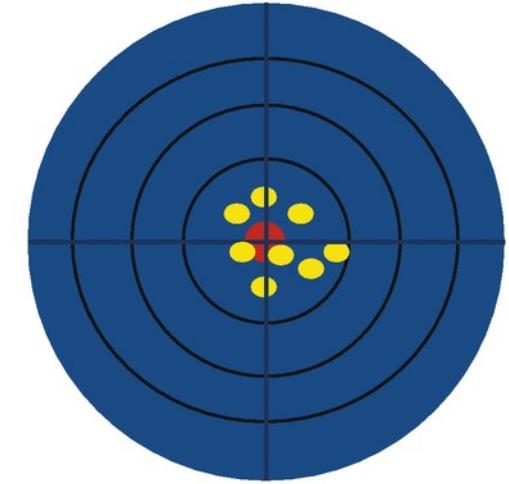
Random Error

Inconsistent estimates
but unbiased



Systematic Error

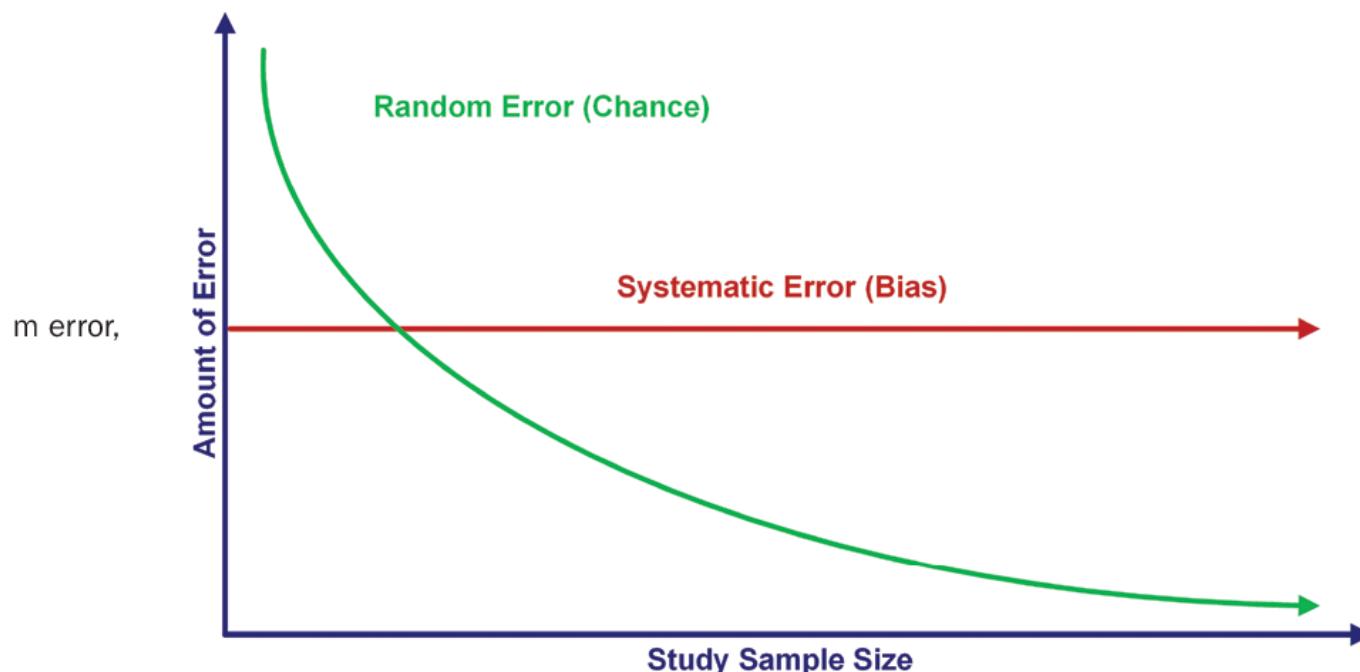
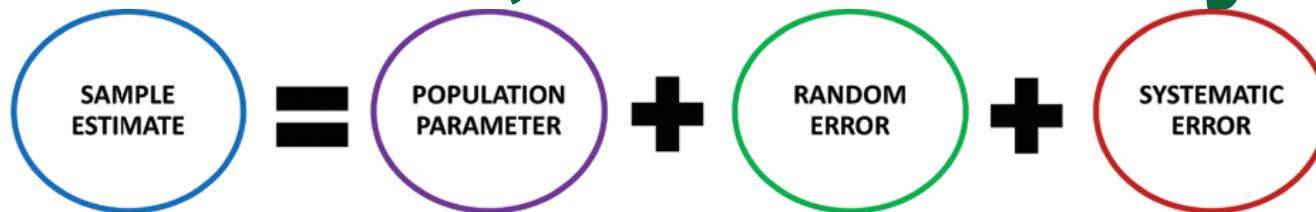
Biased estimates
but consistent



The Truth

Unbiased estimate

Random error may decrease with sample size, but not bias, so bias is systematic



Sources of bias

- Study design, including recruitment and follow-up
- Data collection and quality monitoring of our data
- Analytical procedures

Types of bias

- 1. Selection bias
- 2. Information bias
- 3. Many people describe confounding is also a form of bias, but probably not correct (more in lecture 4 – confounding)

Selection bias

If the selection and/or withdrawal of participants is related to the outcome and exposure, selection bias can occur.



External validity

Internal validity

Target population

The population to which you like to generalize your research

Example
People residing in SK aged 12 and greater

Define a reachable population

Source population

At-risk people who are accessible for recruitment from the target population

Example
SK residents in 2022 Spring EXCLUDING military personnel, incarcerated, living in first nation reserves, homeless, out-of-province healthcare coverage.

Selection bias

Selection bias

Study population or sample

A sample of people included in your study for analysis

Contact people listed in a sampling frame

Example
Participants agreed to be contacted and responded to questionnaires

- Note: terminologies are not consistent in the literature (e.g., source population is often referred as study base by some authors)

Information bias

- Measurement error
- Misclassification
 - Differential or non-differential

Source of information bias

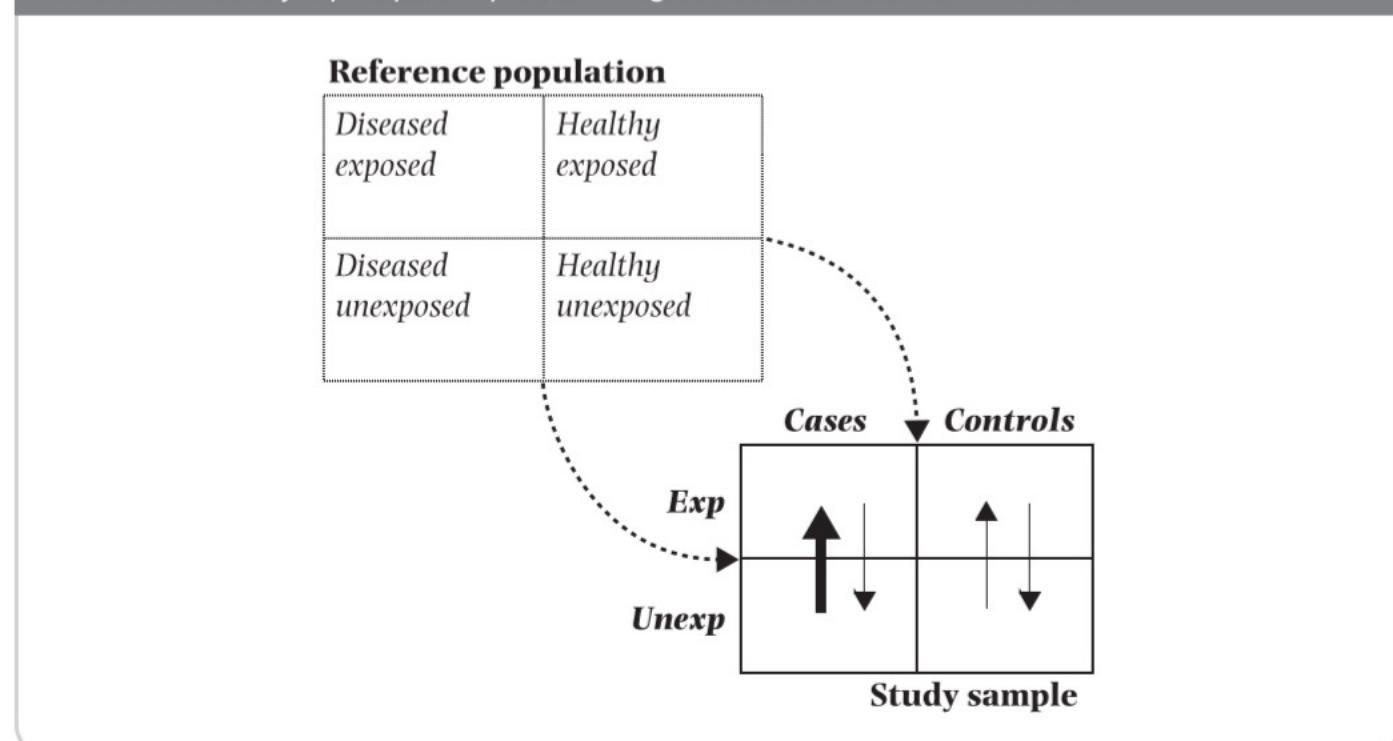
- Imperfect data collection or coding of variables resulting in wrong values of exposure, outcome, and confounding variables
 - Inaccurate diagnosis
 - Mis-reporting of health status by participants (e.g., recall bias)
 - Improper data collection by interviewers (e.g., interviewer bias)

There are 2 types of information bias: Misclassification and Measurement error

Misclassification

- Information bias in categorical (class) variables, whose magnitude of errors measured by
 - a) Sensitivity
 - b) Specificity
 - c) Reliability

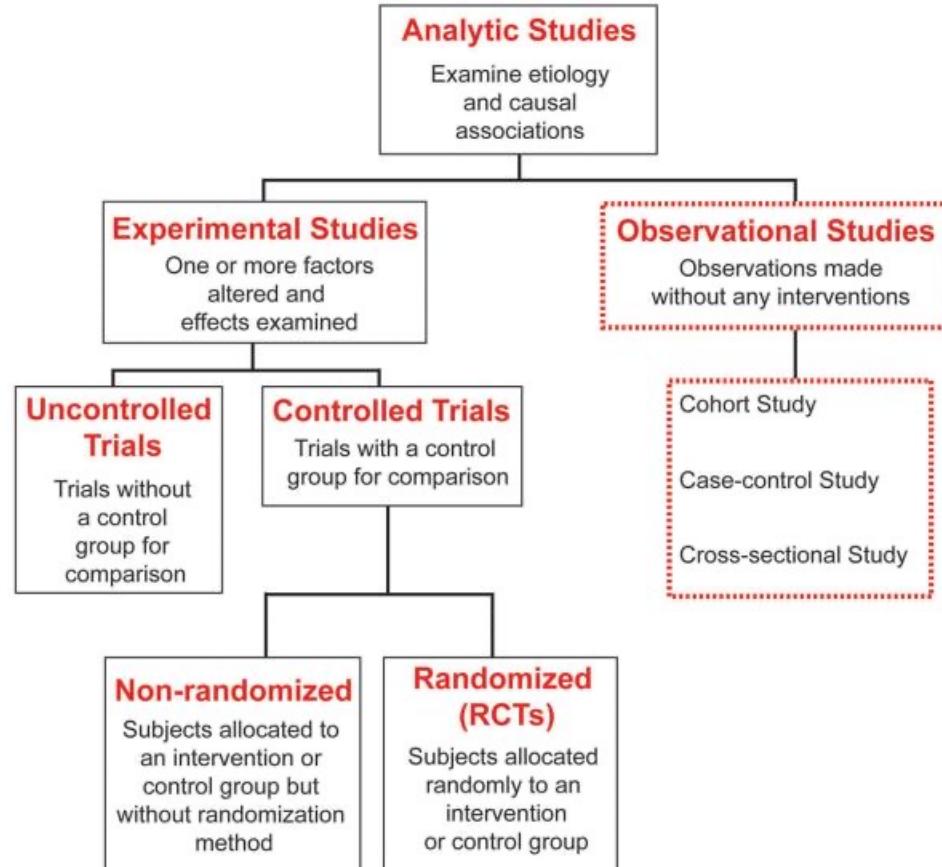
FIGURE 4-3 Misclassification (information) bias: some degree of misclassification of the exposure information exists in both cases and controls, but unexposed cases in this example tend to mistakenly report past exposure to a greater extent than do controls.



Hierarchy of Study Design (mine)

- Systematic review + meta analysis
- Randomized designs (RCT)
- Natural experiments
- Cohort studies
- Case control studies
- Cross-sectional & ecological studies

Analytic Study Designs



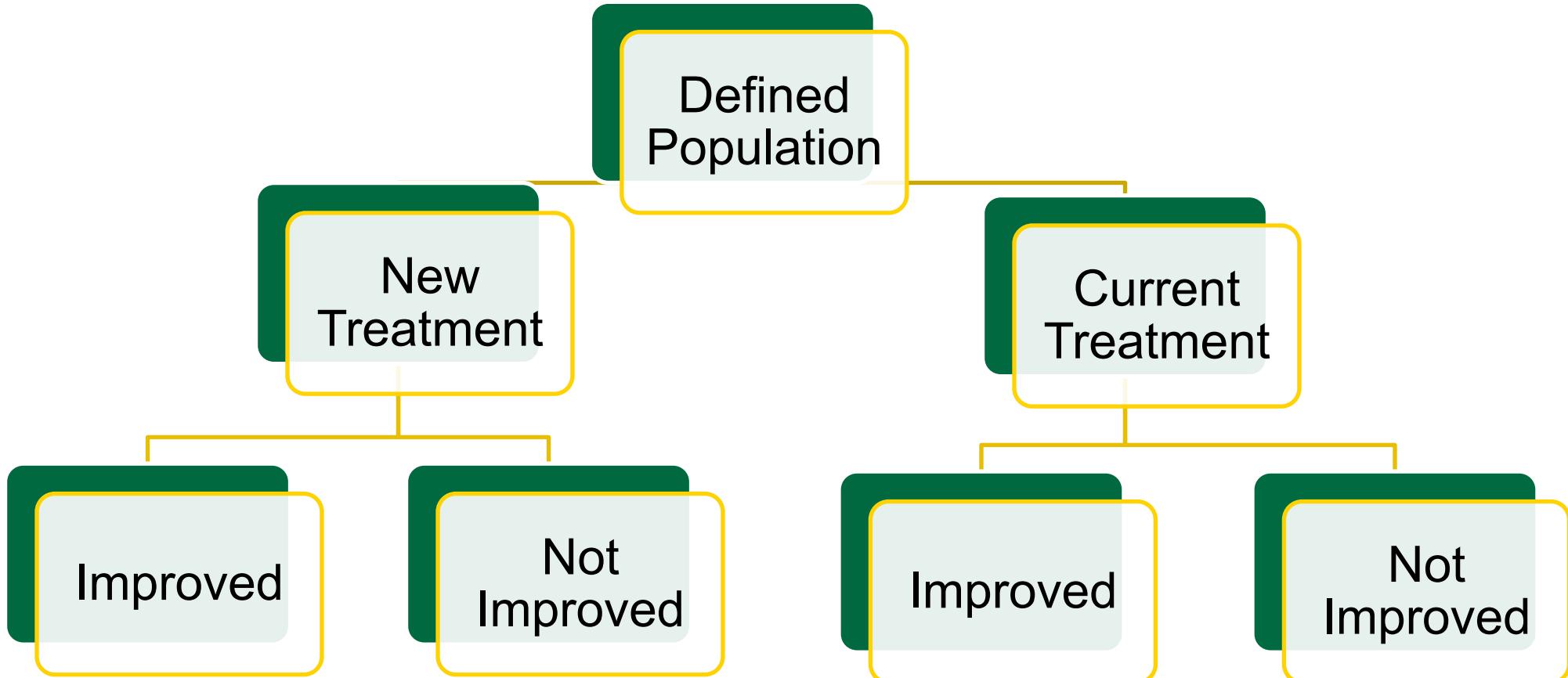
The goal of analytic studies is to identify and evaluate causes or risk factors of diseases or health related events.

From: Plast Reconstr Surg. 2010 Dec; 126(6): 2234–2242.
doi: [10.1097/PRS.0b013e3181f44abc](https://doi.org/10.1097/PRS.0b013e3181f44abc)

Experimental Studies

- Randomized Controlled Trials
- Non-randomized Controlled Trials
 - a) Historical Controls
 - b) Simultaneous Nonrandomized Controls

Randomized Clinical Trials



Randomized Controlled Trials

Four possible conclusions from study results:

1. Treatments do not differ & we correctly recognize this
2. Treatments do not differ and we determine that they do differ
3. The treatments differ but we determine that they do not differ
4. The treatments differ & we correctly recognize this

Possible Outcomes of a Randomized Trial

Decision	Reality	
	Treatments are not different	Treatments are different
Conclude Treatments are not different	Correct Decision	Type II Error (Probability= β)
Conclude Treatments are different	Type I Error (Probability= α)	Correct Decision (Probability = $1-\beta$) (power)

α (Alpha)

- = Probability of making a **Type I Error**
- = Probability of concluding the treatments differ when in reality, they do not differ

β (Beta)

- = Probability of making a **Type II Error**
- = Probability of concluding that the treatments do not differ when in reality, they do differ

Power

- = 1 - Probability of making a Type II Error (β)
- = Probability of Correctly concluding that the treatments differ
- = Probability of detecting a difference between the treatments if the treatments do in fact differ

Estimating Sample Size

Need to Specify:

- 1) Difference in response rates to be detected
- 2) An estimate of the response rate in one of the groups
- 3) Level of statistical significance (alpha)
 - Chosen by investigator. Usually set at 0.05 or 0.01
- 4) Value of desired Power ($1 - \beta$)
 - Chosen by the investigator. Usually set at 80% or 90%.
- 5) Whether test should be one sided or two sided
 - One sided tests require smaller sample sizes and look to see if there is a change in one direction (ie: a drug is beneficial)
 - Two sided tests require larger sample sizes and look to see if there is a change in two directions (ie: drug may be either harmful or beneficial)



Example

Clinical trial of 2 therapies:

- Current therapy cure rate = 40%
- New therapy cure rate anticipated to be 60%
- We set α at 0.05 and power at 80%
- This will be a two-sided test as we want to monitor for both improvements and harm caused by the new therapy.

How many subjects will we need to study?



Estimating Sample Size

Need to Specify:

1) Difference in response rates to be detected **$60\% - 40\% = 20\%$**

2) An estimate of the response rate in one of the groups **40%**

3) Level of statistical significance (α) **0.05**

Choice of level chosen is up to the investigator. Usually set at 0.05 or 0.01

4) Value of desired Power ($1 - \beta$) **80%**

Choice of level chosen is up to the investigator. Usually set at 80% or 90%.

5) Whether test should be one-sided or **two-sided (test for harm & benefit)**

One sided tests require smaller sample sizes and look to see if there is a change in one direction (i.e.: a drug is beneficial)

Two sided tests require larger sample sizes and look to see if there is a change in either direction (i.e.: drug may be either harmful or beneficial)

TABLE 7-4. Number of Patients Needed in Each Group to Detect Various Differences in Cure Rates;
 $\alpha = .05$; Power $(1 - \beta) = .80$ (Two-Sided Test)

Lower of the Two Cure Rates	Differences in Cure Rates Between the Two Treatment Groups													
	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70
.05	420	130	69	44	36	31	23	20	17	14	13	11	10	8
.10	680	195	96	59	41	35	29	23	19	17	13	12	11	8
.15	910	250	120	71	48	39	31	25	20	17	15	12	11	9
.20	1,090	290	135	80	53	42	33	26	22	18	16	12	11	9
.25	1,250	330	150	88	57	44	35	28	22	18	16	12	11	—
.30	1,380	360	160	93	60	44	36	29	22	18	15	12	—	—
.35	1,470	370	170	96	61	44	36	28	22	17	13	—	—	—
.40	1,530	390	175	97	61	44	35	26	20	17	—	—	—	—
.45	1,560	390	175	96	60	42	33	25	19	—	—	—	—	—
.50	1,560	390	170	93	57	40	31	23	—	—	—	—	—	—

Modified from Gehan E: Clinical trials in cancer research. Environ Health Perspect 32:31, 1979.

Crossover in Clinical Trials

Planned Crossover

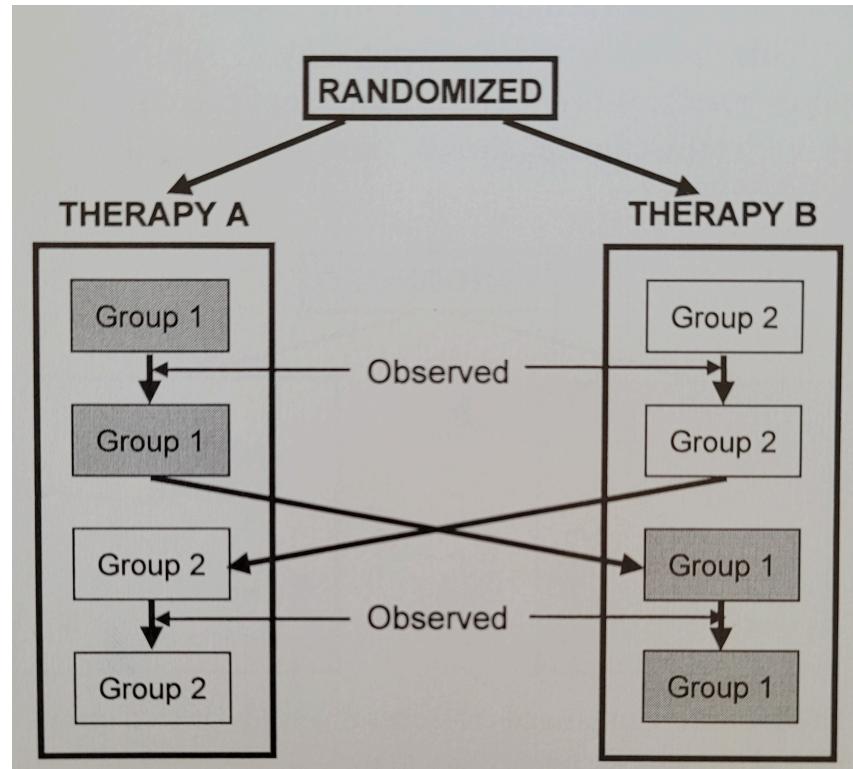
- After a period of observation, subjects assigned to therapy A or B and are switched to other therapy.
- Each subject can serve as his/her own control.
- Keep in Mind:
 - a) Carryover – specify suitable “washout” period is necessary to make sure none of therapy A remains before transfer to therapy B
 - b) Effect of Order – e.g.; diminished enthusiasm regarding treatment over time
 - c) Planned Crossover is not possible if new therapy cures disease

Unplanned Crossover

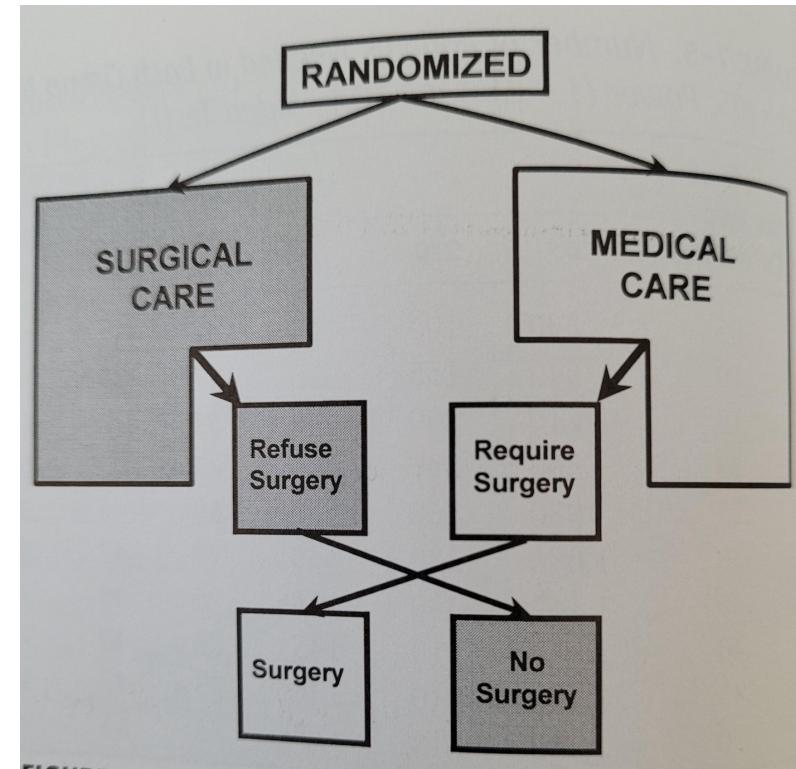
- Randomization was planned but subjects from one group cross-over to the other group
- e.g.; those in Surgical Care Group may change mind about surgery (treatment A) or those in Medical Care Group (treatment B) may require urgent surgery.

Crossover in Clinical Trials

Planned Crossover



Unplanned Crossover



Images From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company

Expressing the Results of Randomized Trials

NNT:

Number of patients who would need to be treated in order to prevent one adverse outcome (e.g.; death)

$$NNT = \frac{1}{(\text{Rate in untreated group}) - (\text{Rate in treated group})}$$

RELATIVE RISK:

The ratio of risk of disease in exposed persons to the risk of disease in nonexposed persons

$$RR = \frac{\text{Risk in exposed}}{\text{Risk in nonexposed}}$$

EFFICACY:

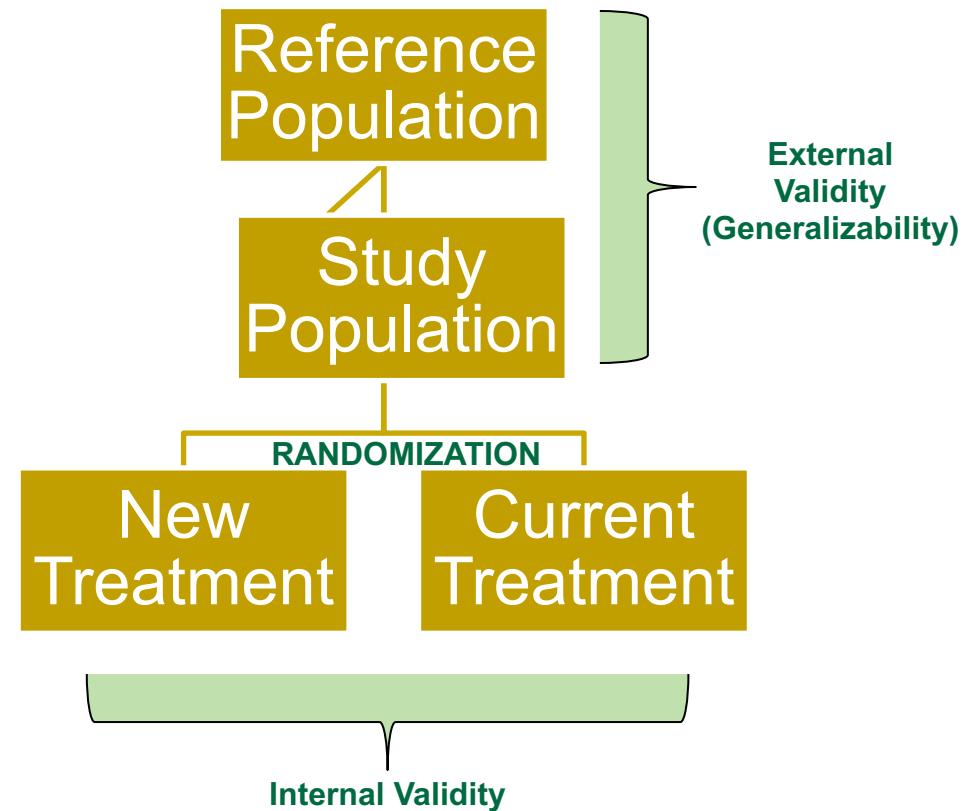
Reduction in risk

$$= \frac{(\text{Rate in those receiving the placebo}) - (\text{Rate in those receiving the Treatment (e.g; vaccine)})}{\text{Rate in those receiving the placebo}}$$

Generalizability of Results of Clinical Trials

Internal Validity: the study is properly done with no major methodical issues

External Validity: the study findings can be generalized to others with the specified disease



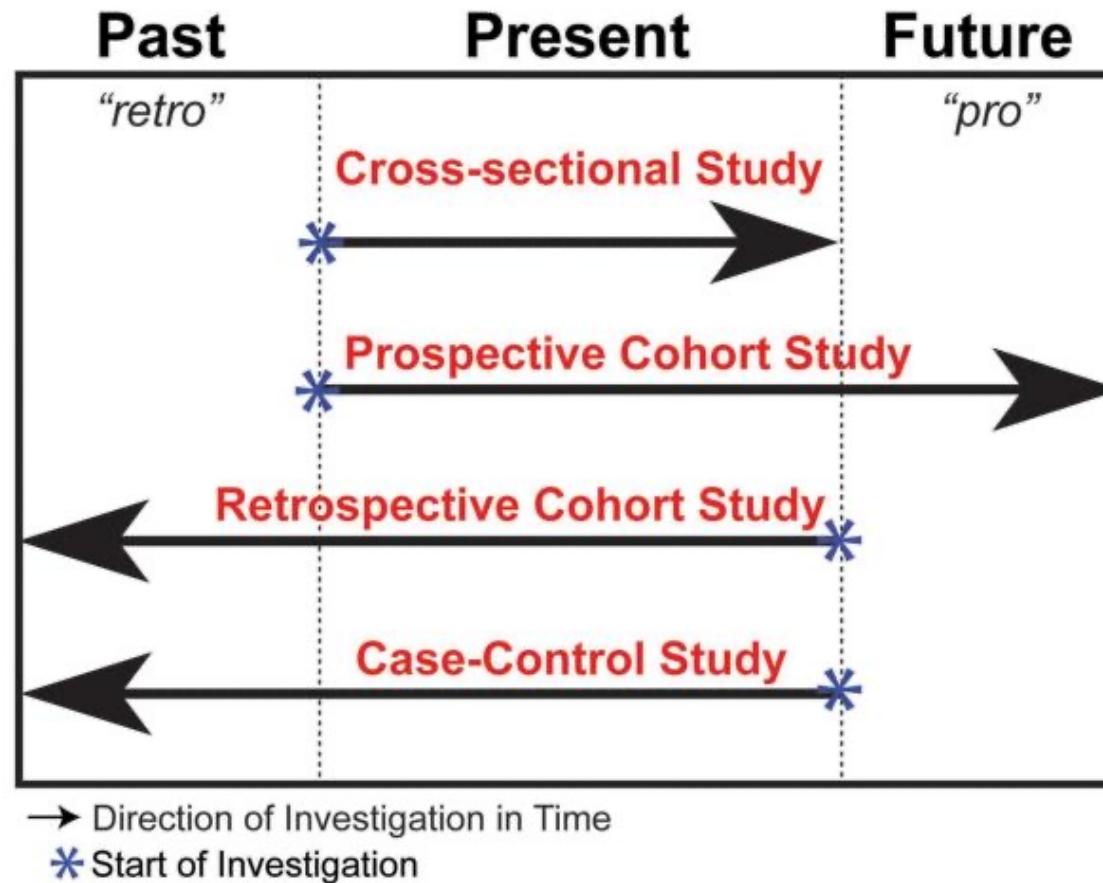
Observational Studies

Cohort Studies

Case-Control Studies

Cross-Sectional Studies

Observational Study Designs



Cohort Study

- Subjects are sorted by exposure status at the start of the investigation
- Both exposed and unexposed groups are selected from the same source population
- Subjects not at risk of developing the disease should be excluded from the study

Two types of Cohort studies:

1. **CONCURRENT** (Concurrent Prospective or Longitudinal Study): from present time into the future
2. **RETROSPECTIVE** (Historical Cohort Studies): look back at past medical events or outcomes (development or no development of disease) for a chosen cohort of subjects in present time

Cohort Study

Advantages

- Can assess sequence of events & causality
- Examine multiple outcomes for an exposure
- Good for investigating rare exposures
- Can calculate rates of disease in exposed and unexposed individuals over time (e.g. incidence, relative risk)

From: Plast Reconstr Surg. 2010 Dec; 126(6): 2234–2242. doi: [10.1097/PRS.0b013e3181f44abc](https://doi.org/10.1097/PRS.0b013e3181f44abc)

Disadvantages

- Require large numbers of subjects to study rare exposures
- Susceptible to 1) bias in assessment of the outcome (can use blinding to correct for bias) 2) analytic bias & 3) selection bias

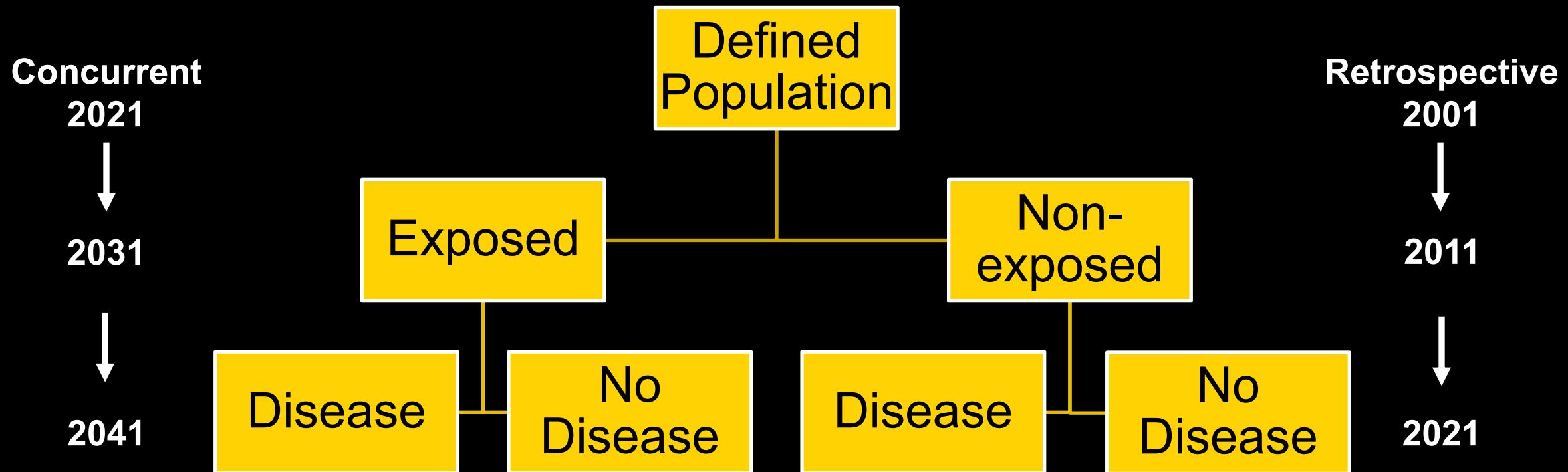
Prospective Cohort Study

- Expensive to conduct
- May require long durations for follow-up
- Maintaining follow-up may be difficult
- Loss to follow-up or withdrawals

Retrospective Cohort Study

- Recall bias or information bias
- Reduced control over variables

Cohort Study



Design of a Cohort Study

Step 2: Follow to see whether disease develops or does not develop

Step 3: Analyze the differences among those exposed and nonexposed

	Disease Develops	Disease Does Not Develop	Totals	Incidence Rates of Disease
Exposed	a	b	$a + b$	$\frac{a}{a + b}$
Nonexposed	c	d	$c + d$	$\frac{c}{c + d}$

Step 1:
First Select {

Case-Control Study

Advantages

- Good for examining rare outcomes or outcomes with long latency
- Relatively quick to conduct
- Relatively inexpensive
- Requires comparatively few subjects
- Existing records can be used
- Multiple exposures or risk factors can be examined

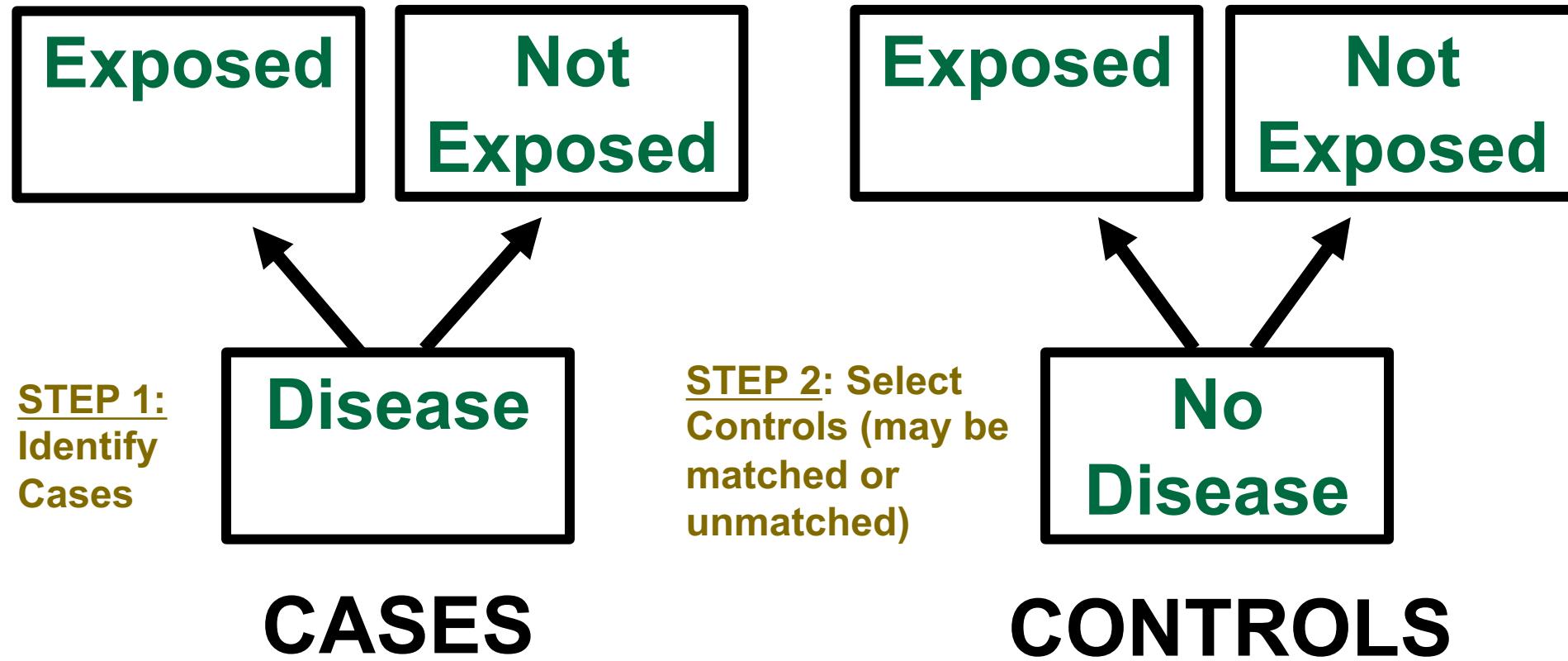
Disadvantages

- Susceptible to recall bias or information bias
- Difficult to validate information
- Control of extraneous variables may be incomplete
- Selection of an appropriate comparison group may be difficult
- Rates of disease in exposed and unexposed individuals cannot be determined

From: Plast Reconstr Surg. 2010 Dec; 126(6): 2234–2242.
doi: [10.1097/PRS.0b013e3181f44abc](https://doi.org/10.1097/PRS.0b013e3181f44abc)

Case-Control Study

STEP 3: Measure exposures or risk factors & compare presence or absence or exposures among cases and controls.



Design of a Case-Control Study

		<u>Step 1:</u> First Select	
		Cases (With Disease)	Controls (without Disease)
<u>Step 2:</u> Then Measure Past Exposure	Were Exposed	a	b
	Were not exposed	c	d
	Total	$a + c$	$b + d$
	Proportions Exposed	$\frac{a}{a + c}$	$\frac{b}{b + d}$

Matching in a Case-Control Study

Matching is used to improve comparability of cases and controls.

Individual Matching: Each case is paired with a control in order to reduce differences of certain demographic variables such as age, sex and race (typical confounders of disease).

Group Matching: Selecting controls so that a proportion of controls with a certain characteristic is the same as the proportion of cases with the same characteristic

Using a Matching Strategy

Advantages

- Eliminate influence of measurable confounders (e.g. age, sex)
- Eliminate influence of confounders that are difficult to measure
- May improve study efficiency (i.e. smaller sample size)

Disadvantages

- May be time-consuming and expensive
- Decision to match and confounding variables to match upon are decided at the outset of the study
- Matched variables cannot be examined in the study
- Requires a matched analysis
- Vulnerable to overmatching: when matching variable has some relationship with the outcome

From: Plast Reconstr Surg. 2010 Dec; 126(6): 2234–2242.
doi: [10.1097/PRS.0b013e3181f44abc](https://doi.org/10.1097/PRS.0b013e3181f44abc)



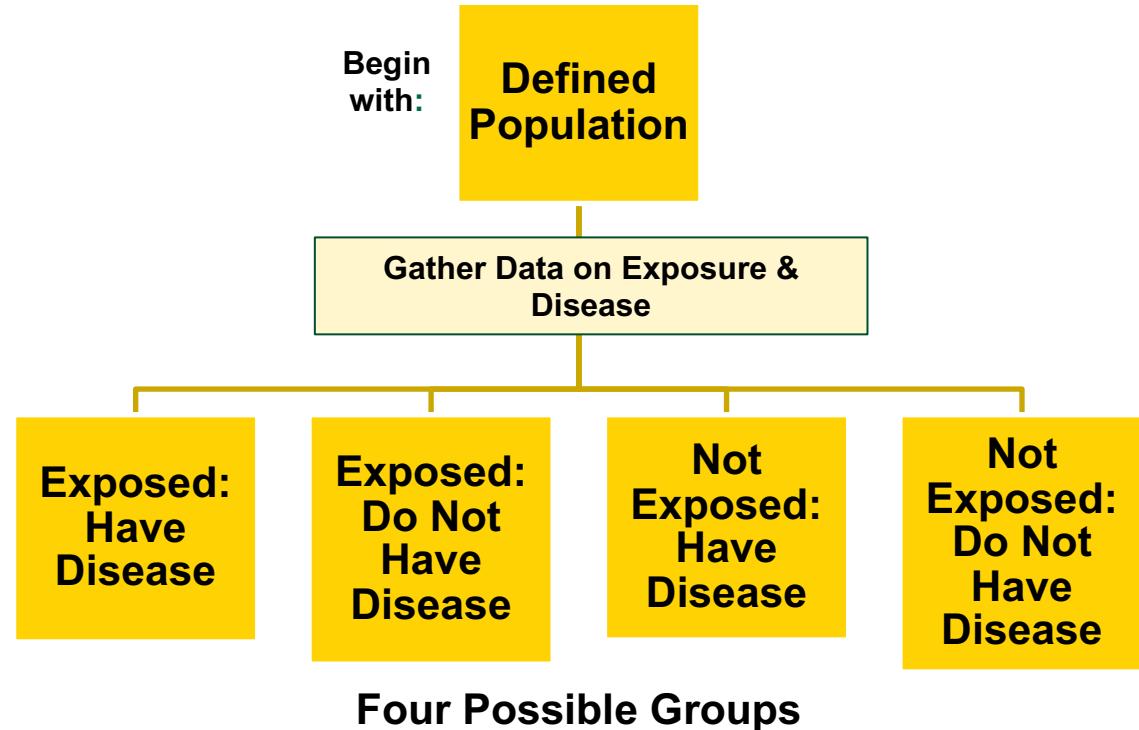
Multiple Controls

Multiple Controls of the Same Type: Having more controls to each case can increase statistical power where there are limited number of cases. (e.g.; hospital controls)

Multiple Controls of Different Types: May be beneficial when the controls may be a subset of nondiseased individuals with a different exposure experience. To address this, an additional control group may be chosen (e.g.; neighbourhood controls)

Cross-Sectional Studies

- Also called Prevalence Studies.
- Both the exposure and the disease outcome are determined simultaneously for each subject at a certain point in time.
- Identified cases are prevalent cases of the disease being studied. We can determine that an individual has been exposed and has the disease but we do not know the duration for which they have had the disease.
- We can calculate Prevalence of Disease & Prevalence of Exposure



Design of a Cross-Sectional Study

	Disease	No Disease
Exposed	a	b
Not Exposed	c	d

Prevalence of DISEASE compared in exposed and nonexposed

$$\frac{a}{a+b} \quad vs \quad \frac{c}{c+d}$$

Prevalence of EXPOSURE compared in diseased and nondiseased

$$\frac{a}{a+c} \quad vs \quad \frac{b}{b+d}$$

Cross-Sectional Study Considerations

- Identifying ***prevalent*** cases *not incident* (new) cases
- Prevalent cases may not be representative of all cases in the population (e.g., would not include those who died of the disease)
- Association of exposure and disease may be due to survival rather than risk of development
- Not possible to establish a temporal relationship as disease and exposure determined at same time and therefore causal relationships cannot be determined.



Estimating Risk

Absolute Risk:

- Defined as the incidence of a disease in a population
- Can show magnitude of risk in a specific group of people with a specific exposure, but does not include consideration of the risk of disease to people who were not exposed

Relative Risk:

- The ratio of the risk of disease in exposed persons to the risk of disease in nonexposed persons



Relative Risk

Relative Risk = Risk in exposed

Risk in nonexposed

If $RR = 1$ Risk in exposed equal to risk in non-exposed (no association)

If $RR > 1$ Risk in exposed greater than risk in non-exposed (positive association; possibly causal)

If $RR < 1$ Risk in exposed less than risk in non-exposed (negative association; possibly protective)

Relative Risk in Cohort Studies

	Disease Develops	Disease Does Not Develop	Totals	Incidence Rates of Disease
Exposed	a	b	$a + b$	$\frac{a}{a + b}$
Nonexposed	c	d	$c + d$	$\frac{c}{c + d}$

Incidence in exposed

Incidence in nonexposed

$$\text{Relative Risk} = \frac{\text{Incidence in exposed}}{\text{Incidence in nonexposed}} = \frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)}$$

Cohort Study of Smoking & Coronary Heart Disease (CHD)

	Disease Develops	Disease Does Not Develop	Totals	Incidence per 1,000 per year
Exposed	a 84	b 2916	a + b 3,000	$\left(\frac{a}{a+b}\right) * 1000 = 28$
Nonexposed	c 87	d 4913	c + d 5,000	$\left(\frac{c}{c+d}\right) * 1000 = 17.4$

$$\text{Relative Risk} = \frac{\text{Incidence in exposed}}{\text{Incidence in nonexposed}} = \frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)} = \frac{28}{17.4} = 1.61$$

The Odds Ratio (Relative Odds)

In order to calculate a relative risk we need to know the incidence of disease in exposed and nonexposed individuals. This can be determined in a cohort study. In a case-control study however, we do not have this information because we start with diseased and non-diseased persons. Therefore, we cannot calculate relative risk directly in a case-control study. ***The odds ratio is the measure of association used in case-control studies.*** Under many conditions, the odds ratio can provide a very good estimate of the relative risk.

$$\text{Odds of an Event} = \frac{\text{Probability of an event occurring}}{\text{Probability of an event not occurring}} = \frac{P}{1-P}$$

Odds Ratio

Odds Ratio in a COHORT STUDY

	Develop Disease	Do Not Develop Disease
Exposed	a	b
Not Exposed	c	d

ODDS RATIO

= Odds that an exposed person develops disease
 Odds that a non-exposed person develops disease

$$= \frac{a/b}{c/d} = \frac{ad}{bc}$$

Odds Ratio in a CASE-CONTROL STUDY

	Cases (With Disease)	Controls (Without Disease)
History of Exposure	a	b
No History of Exposure	c	d

ODDS RATIO

= Odds that a case was exposed
 Odds that a control was exposed

$$= \frac{a/c}{b/d} = \frac{ad}{bc}$$

Interpreting the Odds Ratio

Same interpretation as Relative Risk.

If $OR = 1$ Exposure is not related to disease

If $OR > 1$ Exposure is positively related to disease

If $OR < 1$ Exposure is negatively related to the disease

The Odds Ratio can be a good estimate of Relative Risk when...

1. **Cases** are representative (in respect to history of exposure) of all people with the disease in the population that the cases are drawn from.
2. **Controls** are representative (in respect to history of exposure) of all people without the disease in the population that the cases are drawn.
3. When the disease is **rare**.

Ex: Infrequent Disease RR ≈ OR

	Develop Disease	Do Not Develop Disease	
Exposed	200	9800	10,000
Not Exposed	100	9900	10,000

$$\text{Relative Risk} = \frac{200/10,000}{100/10,000} = 2$$

$$\text{Odds Ratio} = \frac{200 \times 9900}{100 \times 9800} = 2.02$$

Ex: Frequent Disease RR ≠ OR

	Develop Disease	Do Not Develop Disease	
Exposed	50	50	100
Not Exposed	25	75	100

$$\text{Relative Risk} = \frac{50/100}{25/100} = 2 \quad \neq \quad \text{Odds Ratio} = \frac{50 \times 75}{25 \times 50} = 3$$

Attributable Risk

- How much of the disease is due to a specific exposure?
- Tells us the potential for prevention of a disease if the exposure under study is eliminated

Incidence in the **EXPOSED** group = Incidence NOT due to exposure (background incidence)
+ incidence due to the exposure

Incidence in the **NONEXPOSED** group = Incidence NOT due to exposure (background incidence)

Attributable Risk Calculations

Incidence Attributable to Exposure

In Exposed Group

$$\left(\text{Incidence in exposed group} \right) - \left(\text{Incidence in nonexposed group} \right)$$

In Total Population

$$\left(\text{Incidence in total population} \right) - \left(\text{Incidence in nonexposed group} \right)$$

Proportion of Incidence Attributable to Exposure

In Exposed Group

$$\frac{\left(\text{Incidence in exposed group} \right) - \left(\text{Incidence in nonexposed group} \right)}{\text{Incidence in exposed group}}$$

In Total Population

$$\frac{\left(\text{Incidence in total population} \right) - \left(\text{Incidence in nonexposed group} \right)}{\text{Incidence in total population}}$$

Study Designs & Measures of Association

MEASURES OF ASSOCIATION	STUDY DESIGN		
	Case-Control	Cohort	Randomized Controlled Trial
Absolute Risk		YES	YES
Relative Risk		YES	YES
Odds Ratio	YES	YES	YES

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

1. Several studies have found that approximately 85% of cases of lung cancer are due to cigarette smoking. This measure is an example of:
 - a) An Incidence Rate
 - b) An Attributable Risk
 - c) A Relative Risk
 - d) A Prevalence Risk
 - e) A Proportionate Mortality Ratio

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

1. Several studies have found that approximately 85% of cases of lung cancer are due to cigarette smoking. This measure is an example of:
 - a) An Incidence Rate
 - b) An Attributable Risk
 - c) A Relative Risk
 - d) A Prevalence Risk
 - e) A Proportionate Mortality Ratio

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

2. In a study of a disease in which all cases that developed were ascertained, if the relative risk for the association between a factor and the disease is equal to or less than 1.0, then:
 - a) There is no association between the factor and the disease
 - b) The factor protects against development of the disease
 - c) Either matching or randomization has been unsuccessful
 - d) The comparison group used was unsuitable, and a valid comparison is not possible
 - e) There is either no association or a negative association between the factor and the disease

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

2. In a study of a disease in which all cases that developed were ascertained, if the relative risk for the association between a factor and the disease is equal to or less than 1.0, then:
 - a) There is no association between the factor and the disease
 - b) The factor protects against development of the disease
 - c) Either matching or randomization has been unsuccessful
 - d) The comparison group used was unsuitable, and a valid comparison is not possible
 - e) There is either no association or a negative association between the factor and the disease

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

3. In a study begun in 1965, a group of 3,000 adults in Baltimore were asked about alcohol consumption. The occurrence of cases of cancer was studied in the group between 1981 and 1995. This is an example of a:
 - a) Cross-sectional study
 - b) Concurrent cohort study
 - c) Retrospective cohort study
 - d) Clinical Trial
 - e) Case-Control Study

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

3. In a study begun in 1965, a group of 3,000 adults in Baltimore were asked about alcohol consumption. The occurrence of cases of cancer was studied in the group between 1981 and 1995. This is an example of a:
 - a) Cross-sectional study
 - b) Concurrent cohort study
 - c) Retrospective cohort study
 - d) Clinical Trial
 - e) Case-Control Study

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

4. In a case-control study, which of the following is (are) true?
- a) The proportion of cases with the exposure is compared with the proportion of controls with the exposure
 - b) Disease rates are compared for people with the factor of interest and for people without the factor of interest
 - c) The investigator may choose to have multiple comparison groups
 - d) Recall bias is a potential problem
 - e) a, c, and d

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

4. In a case-control study, which of the following is (are) true?
- a) The proportion of cases with the exposure is compared with the proportion of controls with the exposure
 - b) Disease rates are compared for people with the factor of interest and for people without the factor of interest
 - c) The investigator may choose to have multiple comparison groups
 - d) Recall bias is a potential problem
 - e) a, c, and d

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

5. A randomized trial comparing the efficacy of two drugs showed a difference between the two (with a P value < 0.5). Assume that in reality, however, the two drugs do not differ. This therefore an example of:
- a) Type I error (α error)
 - b) Type II error (β error)
 - c) $1 - \alpha$
 - d) $1 - \beta$
 - e) None of the above

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

5. A randomized trial comparing the efficacy of two drugs showed a difference between the two (with a P value < 0.5). Assume that in reality, however, the two drugs do not differ. This therefore an example of:

- a) Type I error (α error)
- b) Type II error (β error)
- c) $1 - \alpha$
- d) $1 - \beta$
- e) None of the above

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

6. In many studies examining the association between estrogens and endometrial cancer of the uterus, a one-sided significance test was used. The underlying assumption justifying a one-sided rather than a two-sided test is:
- a) The distribution of the proportion exposed followed a “normal” pattern
 - b) The expectation before doing the study was that estrogens cause endometrial cancer of the uterus
 - c) The pattern of association could be expressed by a straight line function
 - d) A type II error was the most important potential error to avoid
 - e) Only one control group was being used

REVIEW QUESTIONS

(From: Gordis, L. (2000). Epidemiology: Second Edition. Philadelphia, Pennsylvania: W.B. Saunders Company)

6. In many studies examining the association between estrogens and endometrial cancer of the uterus, a one-sided significance test was used. The underlying assumption justifying a one-sided rather than a two-sided test is:
- a) The distribution of the proportion exposed followed a “normal” pattern
 - b) The expectation before doing the study was that estrogens cause endometrial cancer of the uterus
 - c) The pattern of association could be expressed by a straight line function
 - d) A type II error was the most important potential error to avoid
 - e) Only one control group was being used