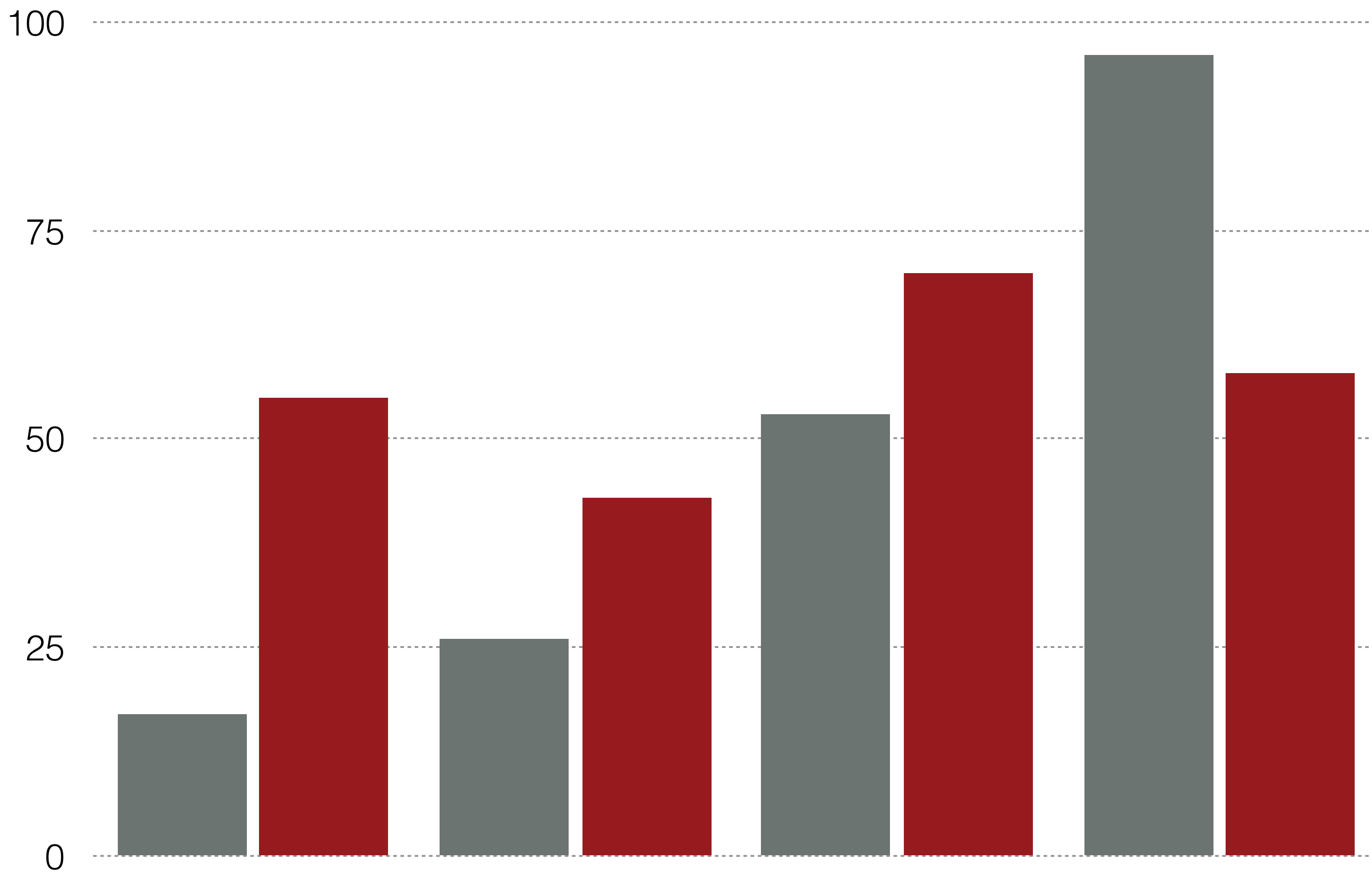
Gender Analysis

Week 3

Objectives

1. Understand how to practically conduct gender analysis

2. Learn and apply measures of frequency and association



Measures of Disease

Frequency and Association

Measures of Disease Frequency

Need measurements that describe the existence or development of disease

Number of cases is useful

If you want to assess if there's a problem, need more information

- Excess vs expected

Measures of Disease Frequency

Need:
frequency of disease expressed
per unit size of the population
in a **specified time period**

$$\frac{\text{number of cases}}{\text{population}}$$ in specified time period

Various measures of frequency, most fall into either prevalence and incidence

Prevalence (P)

- Refers to those with the disease at a given point in time (snapshot)
- =(# of existing cases)/(# in total population) at a point in time
- Interpretation: On DATE, X% of POPULATION in LOCATION had OUTCOME.
- Range of values: 0-1, or 0-100%
- Seen as the probability that an individual will have a disease at a given point in time

Cumulative incidence (CI, "incidence")

- Refers to those who develop a disease among those who are at baseline free of disease and at risk, over a given time period
- =(# of new cases)/(# at risk in population) in a specified time period
- Interpretation: Between DATE1 and DATE2, X% of POPULATION at risk in LOCATION developed OUTCOME.
- Range of values: 0-1, or 0-100%
- Seen as the probability that an individual will have a disease within a specified time period
- Requires everyone followed for the same specified time period, else use incidence rate (IR)

Loss to follow-up

- Refers to subjects with missing portion of follow-up
- May still be at risk and not developed disease
- May have developed disease

Solution: incidence rate (IR)

- Rate new cases of disease occur in a population at risk for the disease
- =(# of new cases developing over study period)/(total person time)
- Range of values: 0-infinity
- Difficult to intuitively understand

Measures of Association

After measures of disease frequency in different groups, goal:

- Summarize into single measure
    - Magnitude (strength) of the relationship (association) between the exposure/risk factors and outcome

        1. Summarize the data in table
        2. Calculate measures of disease frequency for each of the exposure groups
        3. Combine and calculate measures of association

1. Summarize the Data

		Disease (outcome)		
		Yes	No	
Exposure	Yes	a	b	a+b
	No	c	d	c+d
		a+c	b+d	N

a=# exposed and have the disease
b=# exposed and do not have the disease
c=# not exposed and have the disease
d=# both not exposed and do not have the disease

		Brain Cancer		
		Yes	No	
Cell Phone Use	Yes	a	b	a+b
	No	c	d	c+d
		a+c	b+d	N

		Brain Cancer		
		Yes	No	
Cell Phone Use	Yes	a	-	$PT_{exposed}$
	No	c	-	$PT_{unexposed}$
		a+c	-	PT_{total}

2. Calculate Measures of Disease Frequency

Cohort study with individuals in the denominator (count data)

- $CI_{exposed}$=(# exposed cases)/(# exposed)=a/a+b
- $CI_{unexposed}$=(# unexposed cases)/(# unexposed)=c/c+d

Cohort study with person-time denominator (person-time data)

- $IR_{exposed}$=(# exposed cases)/(exposed person-time)=a/$PT_{exposed}$
- $IR_{unexposed}$=(# unexposed cases)/(unexposed person-time)=c/$PT_{unxposed}$

3. Combine and calculate measures of association

Various measures of association

- Divide measures of frequency
  - Ratio measures (relative scale)
- Subtract the two values
  - Difference measures (absolute scale)

Will depend upon the study design

Ratio Measures

Relative Risk: Generic term for risk ratio, rate ratio, or odds ratio

Risk ratio: cumulative incidence ratio

- $CI_{exposed}/CI_{unexposed}$ in cohort study with count data

Rate ratio: incidence rate ratio

- $IR_{exposed}/IR_{unexposed}$ in cohort study with person-time data

Odds ratio (OR)

- OR=ad/bc in case-control study
  - Cannot calculate incidence rates, since there is no rate of development of the disease
  - Individuals are selected because they either do or do not have the disease

Ratio Measures

Risk Ratio~Rate Ratio~Odds Ratio, in meaning, calculated differently

Range of values 0 to infinity, null value = no association: RR = 1.0

- Relative risk >1.0 positive association: exposure is associated with an increased risk of disease
- Relative risk <1.0 inverse association: exposure is associated with a decreased risk of disease

To interpret any of these measures in a succinct sentence:

- The EXPOSED have X times the RISK/RATE/ODDS of having the OUTCOME compared to the UNEXPOSED.

You can also interpret a relative risk using a percentage.

- When the relative risk is <1 and therefore protective: The EXPOSED have [(1.00-X)*100]% DECREASED RISK/RATE/ODDS of having the OUTCOME compared to the UNEXPOSED.

- When the relative risk is >1 and therefore harmful: The EXPOSED have [(1.00-X)*100]% INCREASED RISK/RATE/ODDS of having the OUTCOME compared to the UNEXPOSED.

- When the relative risk is ≥2, it is easier to understand the following: The EXPOSED have (X*100)% the RISK/RATE/ODDS of having the OUTCOME compared to the UNEXPOSED.

Example 1 (when the relative risk is <1): Individuals who wore sunscreen on Crane Beach in July 2014 have 0.6 times the odds (or 40% decreased odds) of developing a sunburn compared to individual who didn't wear any sunscreen.

Example 2 (when the relative risk is >1 but <2): Adults who ate 2 portions of red meat have 1.7 times the risk (or 70% increased risk) of having a myocardial infarction compared to women you ate 1 portion of red meat.

Example 3 (when the relative risk is ≥2): Women who used talc powder have 2.2 times the risk (or 222% the risk) of having the ovarian cancer compared to women you never used talc powder.

Always include as much information as you have available to you (e.g., dates, exposure definitions, geographic location).

Risk Ratio Example

		Lung Cancer		
		Yes	No	
Smoking	Yes	36	39,964	40,000
	No	4	59,996	60,000
		40	99,960	100,000

$CI_{exposed}$=a/a+b=36/40,000=9/10,000

$CI_{unexposed}$=c/c+d=4/60,000=0.7/10,000

Risk ratio=$CI_{exposed}/CI_{unexposed}$

=(36/40,000)/(4/60,000)

=13.5

Women who are smokers had 13.5 times the risk of lung cancer as nonsmokers (or had a 1350% increased risk)

Rate Ratio Example

		Breast Cancer		
		Yes	No	
Age at 1st birth	35+	49	-	26,757 person-years
	<35	782	-	841,827 person-years
		831	-	

$IR_{exposed}$=a/$PT_{exposed}$=49/26,757=183/100,000 person-years

$IR_{unexposed}$=c/$PT_{unxposed}$=782/841,827=93/100,000 person-years

Rate ratio=$IR_{exposed}$/$IR_{unexposed}$

=(49/26,757)/(782/841,827)

=1.97

Women who had their first birth after age 35 had 1.97 times the rate of breast cancer as those who were <35 years of age at first birth (or had a 97% increased rate).

Odds Ratio Example

		Myocardial Infarction		
		Yes	No	
Current OC Use	Yes	23	304	
	No	133	2616	
		156	3120	3276

Odds ratio=OR

=ad/bc

=23 (2816)/133(304)

=1.6

Women who are current OC users had 1.6 times the odds of a myocardial infarction as nonusers (or had a 60% increased odds)

Difference Measures

Risk difference: General term indicates amount of disease in the exposed group that can be considered due to (attributable to) the exposure, by subtracting out the rate (risk) of disease in the nunexposed group

Assumes causality has been established

Null is 0 (rather than 1 like in ratio measures)

Risk difference

- $CI_{exposed}$-$CI_{unexposed}$ in cohort study with count data

Rate difference

- $IR_{exposed}$-$IR_{unexposed}$ in cohort studies with person-time denominators

Can change to percent ("attributable risk percent")

- $(CI_{exposed}-CI_{unexposed})/CI_{exposed}$
- (OR-1)/OR

Risk Difference Example

		Lung Cancer		
		Yes	No	
Smoking	Yes	36	39,964	40,000
	No	4	59,996	60,000
		40	99,960	100,000

$CI_{exposed}$=a/a+b=36/40,000=9/10,000

$CI_{unexposed}$=c/c+d=4/60,000=0.7/10,000

Risk ratio=$CI_{exposed}$-$CI_{unexposed}$

=(36/40,000)-(4/60,000)

=8.3/10,000

Assuming smoking causes lung cancer, 8.3 cases of lung cancer per 10,000 smokers (exposed) is due to smoking, or could be eliminated among smokers if smoking were eliminated.

Risk Difference Example

		Lung Cancer		
		Yes	No	
Smoking	Yes	36	39,964	40,000
	No	4	59,996	60,000
		40	99,960	100,000

Attributable risk percent=$(CI_{exposed}-CI_{unexposed})/CI_{exposed}$

$=[(36/40{,}000)-(4/60{,}000)]/(36/40{,}000)$

$=92.6\%$

If smoking causes lung cancer, 93% of the lung cancer among smokers is due to smoking, or could be avoided among smokers if smoking were eliminated.

Measures of Frequency and Association

Exercise

For each description below:

- Indicate which measure of disease frequency (prevalence, cumulative incidence, or incidence rate) best describes each of the following.
- Indicate the reason why.

Percentage of freshman girls who become pregnant over the course of their high school years.

- Cumulative incidence: development, uniform follow-up

The percentage of senior boys who are fathers at the time of graduation.

- Prevalence: existing at a point in time

The proportion of women who experience depression on the third postpartum day.

- Prevalence: existing at a point in time

The percentage of infants with congenital heart defects at birth.

- Prevalence: existing at a point in time

The percentage of workers who were all followed for 10 years after beginning employment in a coal mine, who developed lung cancer during that period.

- Cumulative incidence: development, uniform follow-up

The number of first myocardial infarctions (heart attacks) occurring in middle-aged men in 10,000 person-years of observation.

- Incidence rate: development, variable follow-up

The percentage of potential army recruits rejected at their initial enrollment physical exam because of poor vision.

- Prevalence: existing at a point in time

A study was conducted among 400 Canadian women diagnosed with breast cancer.

The goal of the study was to assess whether the type of hospital at which the woman was treated (provincial versus teaching) was related to her survival from the disease over a 10-year period.

The table below shows the outcome variable (the number of women surviving as of March 1999), the exposure variable (type of hospital [provincial or teaching] at which the woman was treated), as well as the extent of her disease at diagnosis.

Extent of Breast Cancer	Type of Hospital Where Treated					
	Provincial Hospital			Teaching Hospital		
	N treated	*%*	*N survived*	*N treated*	*%*	*N survived*
Local	240	80	144	30	30	21
Regional	30	10	9	40	40	20
Metastatic	30	10	3	30	30	6

What is the study design?

- Cohort study
  - Exposure = type of hospital where treated
  - Outcome = survival

Set up a 2x2 table with total breast cancer deaths as the outcome of interest and hospital type (provincial or teaching) as the exposure of interest.

Hospital Type	**Died**		
	Yes	No	
Provincial	144	156	300
Teaching	53	47	100
	197	203	400

Calculate the cumulative incidence ratio (risk ratio) of total breast cancer deaths in the provincial hospitals relative to the teaching hospitals.

Hospital Type	Died		
	Yes	No	
Provincial	144	156	300
Teaching	53	47	100
	197	203	400

$$IR = \frac{(144/300)}{(53/100)} = 0.9$$

Interpret this value in words: Women <u>treated in a provincial hospital</u> had 90% of the risk of <u>dying</u> (or 10% less risk of dying) from breast cancer over 10 years as those <u>treated in a teaching hospital</u>.

Calculate the attributable risk percent among the exposed (those treated in provincial hospitals).

$$AR_E\% = \frac{(R_1 - R_0)}{R_1} * 100 = \frac{(144/300 - 53/100)}{(144/300)} = -10\%$$

Hospital Type	Died		
	Yes	No	
Provincial	144	156	300
Teaching	53	47	100
	197	203	400

Interpret this value in words: Assuming that being treated at a provincial hospital is causally related to a reduction in mortality, 10% of deaths from breast cancer among those treated at the provincial hospital were prevented because they were treated at the provincial hospital.

Using data provided in the original table, suggest an alternative explanation as to why this reduction in mortality could have been observed, besides that being treated at a provincial hospital is causally related to greater survival.

- Confounding by extent of disease at diagnosis
  - 80% of those treated at the provincial hospitals had local disease at diagnosis
  - As compared with only 30% at the teaching hospitals, and local disease has a better survival rate than, for example, metastatic disease

To evaluate possible risk factors for breast cancer, investigators conducted a case-control study

Cases: 1,000 women with breast cancer who had previously given birth using a tumor registry list from Massachusetts

Controls: 1,000 women without breast cancer who had previously given birth were selected at random from voter registration lists

All subjects were interviewed on a variety of exposures including reproductive characteristics, demographic information, and past medical history

525 of those with breast cancer reported having an age at first birth of 35 years or younger

635 women without breast cancer reported having their age at first birth as 35 years or younger

Construct a 2x2 table that summarizes these data on the association between age at first birth of 35 years or younger versus older than 35 years, and breast cancer.

Age at first birth (years)	Breast Cancer		
	Yes	No	
≤ 35	525	635	1160
> 35	475	365	840
	1000	1000	2000

What measure of association would you use to describe the relationship between age at first birth and breast cancer?

- Odds ratio

Calculate that measure of association for this study.

Age at first birth (years)	Breast Cancer		
	Yes	No	
≤ 35	525	635	1160
> 35	475 No	365	840
	1000 635	1000	2000

$$OR = \frac{(525*365)}{(475*635)} = 0.64$$

Interpret the meaning of this value in words.

- Women who have an <u>early age at first birth</u> have 0.64 times the odds (or 64% of the odds, or 36% less odds) of developing <u>breast cancer</u> compared with those who have a <u>late age at first birth</u>.

WGS.151 Gender, Health, and Society
Spring 2016