



Brigham and Women's Hospital

Founding Member, Mass General Brigham

BIOSTATISTICS

Pamela M. Rist, ScD

Associate Epidemiologist

Division of Preventive Medicine, Department of Medicine

Brigham and Women's Hospital

Associate Professor

Harvard Medical School



Pamela M. Rist, ScD



Harvard T.H. Chan School of Public Health
(HSPH), MS and ScD in Epidemiology

Post-doctoral training @ BWH and HSPH

Senior Epidemiologist @ BWH

Associate Professor @ HMS and HSPH

Associate Director, Master of Science in
Epidemiology @ HSPH

Research focus: Epidemiology of stroke,
migraine, and cognitive decline

DISCLOSURES

Dr. Rist has received study pills and pills bottles from Mars Edge for use in an NIH-funded, investigator-initiated pilot study.



OBJECTIVES

To define and calculate the characteristics of a screening test

To calculate and interpret measures of disease frequency and measures of association

To identify and describe the key features of epidemiologic study designs

To define and identify confounding and bias in various study designs

To define key statistical principles including statistical significance, power, and methods of analysis based on variable types



Screening

100 women over the age of 50 received mammograms at a mobile breast cancer screening unit

27 women had findings suspicious for malignancy on the mammogram: 19 of these women were confirmed as having breast cancer by biopsy (true positives)

1 woman had a negative mammogram but in the subsequent year developed breast cancer and is assumed to have had the disease at the time of screening (false negative)

Questions of Interest:

What is the sensitivity and the specificity of the screening test?

What are the positive and negative predictive values?



Sensitivity and Specificity

		Disease Status		
		+	-	
Screening Test	+	19	8	27
	-	1	72	73
		20	80	100

$$\text{Sensitivity} = \frac{T+}{Dx+} = \frac{19}{20} = 95\%$$

$$\text{Specificity} = \frac{T-}{Dx-} = \frac{72}{80} = 90\%$$



Positive and Negative Predictive Values

		Disease Status		
		+	-	
Screening Test	+	19	8	27
	-	1	72	73
		20	80	

$$\text{Positive Predictive Value} = \frac{Dx+}{T+} = \frac{19}{27} = 70\%$$


$$\text{Negative Predictive Value (-)} = \frac{Dx-}{T-} = \frac{72}{73} = 99\%$$

Vignette

In a Mantoux tuberculosis screening program in a high-risk population, the criterion for a positive test was defined as 10 mm of induration. If guidelines change, and the criterion for a positive test is now defined lower, as only 5 mm of induration, which of the following will be true? (Select all that apply.)

- A. Sensitivity will increase
- B. Specificity will decrease
- C. Positive predictive value will increase
- D. False positives will increase
- E. False negatives will decrease

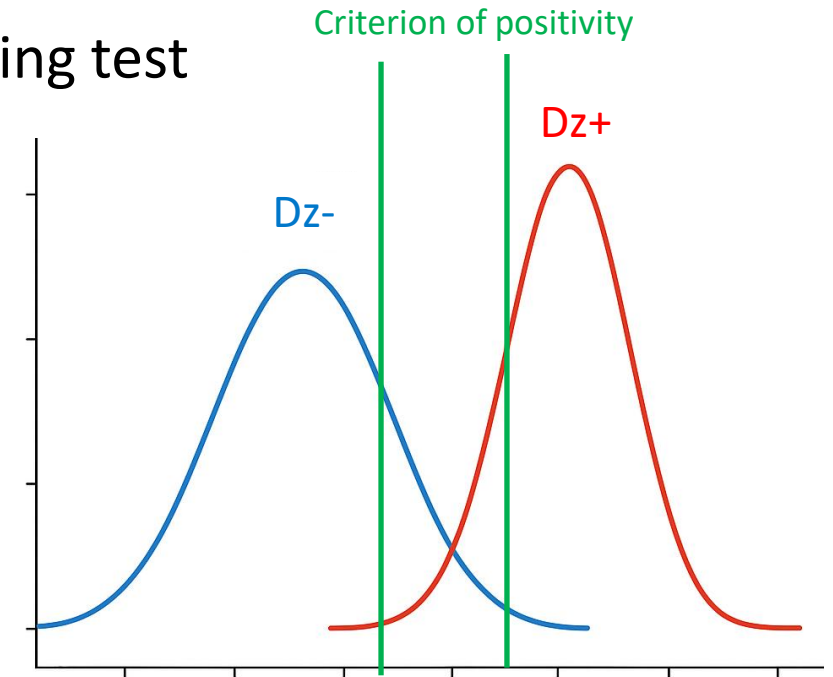


Criterion of Positivity

- Test value on continuum at which the screening test is considered positive
- Influences both sensitivity and specificity of the screening test

↓ Criterion of positivity → ↑ sensitivity
↓ specificity

↑ Criterion of positivity → ↓ sensitivity
↑ specificity



- Sensitivity and specificity are characteristics of the screening test, but can be modified and traded off against each other



Question

In a Mantoux tuberculosis screening program in a high-risk population, the criterion for a positive test was defined as 10 mm of induration. If guidelines change, and the criterion for a positive test is now defined lower, as only 5 mm of induration, which of the following will be true? (Select all that apply.)

- A. Sensitivity will increase**
- B. Specificity will decrease**
- C. Positive predictive value will increase
- D. False positives will increase**
- E. False negatives will decrease**



Vignette

Randomized trial to evaluate effectiveness of a new colon cancer screening program

Among those whose cancers were detected by screening program, average age at diagnosis was 54 years and average age at death was 60 years

- Average survival from diagnosis to death was 6 years

For those detected by clinical symptoms, average age at diagnosis was 56 years and average age at death was 60 years

- Average survival from diagnosis to death was 4 years

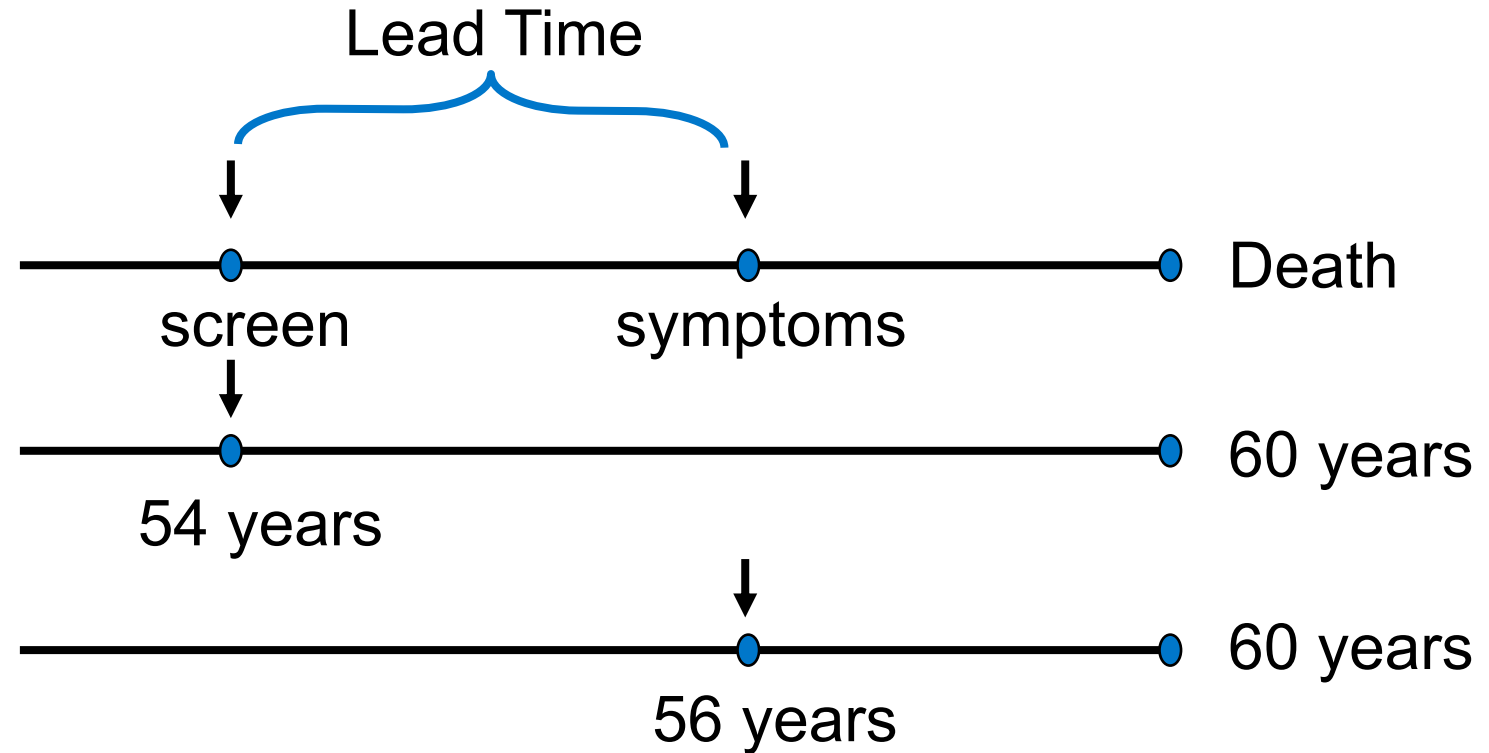
Investigators reported a statistically significant two year increase in survival from colon cancer associated with screening.

What is this an example of?

- A. This is a correct conclusion
- B. Lead time bias
- C. Length time bias



Lead Time Bias



Incorrect Conclusion: Two-year increase in survival associated with screening

Problem: Lead time bias

Solution: Compare age-specific mortality rates



Vignette

Randomized trial to evaluate effectiveness of new colon cancer screening program

Among those whose cancers were detected by screening program, average age at diagnosis was 54 years and average age at death was 60 years

- Average survival from diagnosis to death was 6 years

For those detected by clinical symptoms, average age at diagnosis was 56 years and average age at death was 60 years

- Average survival from diagnosis to death was 4 years

Investigators reported a statistically significant two year increase in survival from colon cancer associated with screening.

What is this an example of?

- A. This is a correct conclusion
- B. Lead time bias**
- C. Length time bias



Other Screening Concepts – Quick Hits

Length bias

- Screening tends to detect cases with a long preclinical course, which should have a longer clinical course and better survival
- May make a screening program appear to have a beneficial effect on survival because people who are destined to have a favorable course are selectively identified

Positive Predictive Value and the Prevalence of Disease

- For rare diseases, a higher prevalence in the screened population will result in a higher positive predictive value
- Often screen high-risk groups to increase the PPV (higher yield for the screening program)



Measures of Disease Frequency

If you see that the prevalence of AIDS in the United States in 2005 is greater than in 1995 (pre-HAART – highly active antiretroviral therapy), what could this be due to?

- A. A change in the incidence rate
- B. A change in the duration of disease
- C. Either or both A and B



Measures of Disease Frequency

$$\text{Prevalence} = \frac{\text{number of existing cases at a point in time}}{\text{total population}}$$

$$\text{Incidence} = \frac{\text{number of new cases during a period of time}}{\text{population at risk}}$$

The proportion of the population that exist with a disease at a point in time (prevalence) depends on both the rate of development of the disease in the population (incidence) as well as the duration of disease from onset to termination (death or cure)

$P \sim \text{Incidence Rate} * \text{Duration}$ (under steady state – no epidemics, no breakthrough treatment)



Measures of Disease Frequency

If you see that the prevalence of AIDS in the United States in 2005 is greater than in 1995 (pre-HAART – highly active retroviral therapy), what could this be due to?

- A. A change in the incidence rate
- B. A change in the duration of disease
- C. **Either or both A and B**



Measures of Association

The table below gives the annual mortality rates from lung cancer and coronary heart disease among cigarette smokers and nonsmokers in a cohort study of British male physicians.

Annual Mortality Rates per 100,000		
	Lung cancer	Coronary heart disease
Smokers	140	669
Non-smokers	10	413

Questions to consider:

Based on these data is smoking a stronger risk factor for lung cancer or CHD?

If smoking were eliminated, could more deaths be eliminated among smokers from lung cancer or CHD?



Measures of Association

Relative risk (RR)

- Measure of the strength of the association between an exposure and outcome
- Formula: $RR = I_e / I_o$
- RR values
 - >1 suggests an increased risk for the exposed group
 - $=1$ suggests no association
 - <1 suggests a decreased risk for the exposed group

Attributable rate/risk (AR) (also known as the risk/rate difference)

- Measure of the public health impact among the exposed
- Formula: $AR_e = I_e - I_o$
- AR_e values
 - If no association, $AR_e = 0$



Measures of Association

Annual Mortality Rates per 100,000		
	Lung cancer	Coronary heart disease
Smokers	140	669
Non-smokers	10	413

$$\begin{aligned}\text{RR (lung cancer)} &= I_e/I_o \\ &= (140/100,000)/(10/100,000) = 14\end{aligned}$$

Those who smoke have 14 times the rate of dying from lung cancer compared to nonsmokers

$$\begin{aligned}\text{RR (CHD)} &= I_e/I_o \\ &= (669/100,000)/(413/100,000) = 1.6\end{aligned}$$

Those who smoke have 1.6 times the rate (or a 60% increased rate) of dying from lung cancer compared to nonsmokers



Measures of Association

Annual Mortality Rates per 100,000		
	Lung cancer	Coronary heart disease
Smokers	140	669
Non-smokers	10	413

$$\begin{aligned}AR_e(\text{lung cancer}) &= I_e - I_0 \\&= (140/100,000) - (10/100,000) \\&= 130/100,000\end{aligned}$$

There are 130 excess deaths per 100,000 smokers; or more broadly for every 100,000 smokers, 130 lung cancer deaths are due to their smoking

$$\begin{aligned}AR_e(\text{CHD}) &= I_e - I_0 \\&= (669/100,000) - (413/100,000) \\&= 256/100,000\end{aligned}$$

There are 256 excess deaths per 100,000 smokers; or more broadly for every 100,000 smokers, 256 CHD deaths are due to their smoking



Measures of association

Based on these data, is smoking a stronger risk factor for lung cancer or CHD?

Lung cancer, since the RR for lung cancer mortality (RR=14) is greater than for CHD (RR=1.6)

But if smoking were eliminated, could more deaths be eliminated among smokers from lung cancer or CHD?

CHD since the AR_e for CHD (256/100,000 smokers) is greater than for lung cancer (130/100,000 smokers)

This reflects that the baseline risk of CHD mortality (413/100,000) is much higher than lung cancer mortality (10/100,000).



Number Needed to Treat

	Nonfatal MI or Death from CHD		
Pravastatin	212	1869	2081
Placebo	274	1804	2078
	486	3673	4159

Number needed to treat = $1/AR_e$

$$= 1 / (212/2081 - 274/2078)$$

$$= 1/0.03$$

$$= 33$$

We would need to treat 33 patients over 5 years (median duration of follow-up in the trial) to prevent one nonfatal MI or CHD death



Measures of Association

- Relative risk and attributable risk/rate can be calculated in randomized clinical trials or in cohort studies
- In these studies, we follow individuals through time for the development of the outcome
- In contrast, for case-control studies, we start by enrolling those with and without the disease
- No incidence rate
- Estimate the relative risk with the odds ratio
- Odds Ratio - estimates the odds of exposure among patients with a disease compared with the odds of exposure among patients without a disease
- $OR = ((A/C)/(B/D)) = AD/BC$

		Disease Status	
		Yes	No
Exposure status	Yes	A	B
	No	C	D



Vignette – Types of Epidemiologic Studies

A study was undertaken to evaluate the relationship between maternal smoking during pregnancy and low birth weight. A total of 350 mothers of low birth weight newborns and 400 mothers of normal weight newborns were interviewed. Among the mothers of low birth weight newborns, 200 reported smoking during the pregnancy, while 200 of the mothers of the normal weight newborns also reported such a history.

Question:

What type of study is this?

- A. Case series**
- B. Cross-sectional study**
- C. Case-control study**
- D. Cohort study**
- E. Randomized clinical trial**



Types of epidemiologic studies

Cross-sectional Survey

Exposure, Outcome

Case Control

Exp

D

Cohort

Exp

D

Clinical Trial (investigator allocates)

Exp

D

Study Type	Advantages	Disadvantages
Cross-sectional	Estimates prevalence	Limited ability to test hypotheses
Case-control	Efficient for rare diseases; study multiple exposures	Susceptibility to bias (selection and recall)
Cohort	Efficient for rare exposures; study several outcomes	Need to minimize loss to follow-up; time consuming and expensive
Clinical trial	Randomization (control of confounding); selection of exposure	Ethical issues; feasibility; cost



Vignette

A study was undertaken to evaluate the relationship between maternal smoking during pregnancy and low birth weight. A total of 350 mothers of low birth weight newborns and 400 mothers of normal weight newborns were interviewed. Among the mothers of low birth weight newborns, 200 reported smoking during the pregnancy, while 200 of the mothers of the normal weight newborns also reported such a history.

Question:

What type of study is this?

- A. Case series
- B. Cross-sectional study
- C. Case-control study**
- D. Cohort study
- E. Randomized clinical trial



Interpretation of Epidemiologic Studies

A study was undertaken to evaluate the relationship between maternal smoking during pregnancy and low birth weight. A total of 350 mothers of low birth weight newborns and 400 mothers of normal weight newborns were interviewed. Among the mothers of low birth weight newborns, 200 reported smoking during the pregnancy, while 200 of the mothers of the normal weight newborns also reported such a history.

Question:

What is the magnitude of the association between smoking and birth weight?



Case-control studies – Measures of Association

	Birth Weight	
	low	normal
Smoking	200	200
No smoking	150	200
	350	400

$$\text{Odds Ratio} = AD/BC = (200*200)/(200*150) = 1.3$$

Mothers who smoked during pregnancy had 1.3 times the risk, or 30% increased risk, of having a low birth weight newborn.



Interpretation of Epidemiologic Studies

In the same case-control study, it is suggested that smoking mothers of low birth weight newborns would tend to deny such an activity due to feelings of guilt, compared to non-smoking mothers of low birth weight newborns, or smoking mothers of normal weight newborns.

This differential reporting would be an example of:

- A. Confounding**
- B. Bias**



Interpretation of Epidemiologic Studies

- Bias – any source of systematic error in the determination of the exposure-outcome association
 - Many different types and names
 - Two main categories:
 - **Selection bias:** bias due to how subjects are selected into study (or choose to participate)
 - **Information or observation bias:** bias due to how information is collected on subjects in study; may result when there is a different level of accuracy or completeness of information between the study groups (i.e. recall bias)
- Confounding – mixture of effect between the association under study and a third variable, associated with exposure and a risk factor for the outcome, which may be responsible in part or totally for the association seen



Interpretation of Epidemiologic Studies

In the same case-control study, it is suggested that smoking mothers of low birth weight newborns would tend to deny such an activity due to feelings of guilt, compared to non-smoking mothers of low birth weight newborns, or smoking mothers of normal weight newborns.

This differential reporting would be an example of:

- A. Confounding
- B. Bias – those with the outcome tend to recall/report exposures differently than those without; this specific type of bias is often termed “recall bias”**



Interpretation of Epidemiologic Studies

In the same case-control study, it is suggested that smoking mothers of low birth weight newborns would tend to deny such an activity due to feelings of guilt, compared to non-smoking mothers of low birth weight newborns, or smoking mothers of normal weight newborns.

This scenario would result in which of the following:

- A. Underestimate of the true relative risk (bias towards the null)**
- B. Overestimate of the true relative risk (bias away from the null)**
- C. The same as the true relative risk**



Interpretation of Epidemiologic Studies

In the same case-control study, it is suggested that smoking mothers of low birth weight newborns would tend to deny such an activity due to feelings of guilt, compared to non-smoking mothers of low birth weight newborns, or smoking mothers of normal weight newborns.

This scenario would result in which of the following:

- A. Underestimate of the true relative risk (bias towards the null) – smoking will not appear as harmful as it actually is**
- B. Overestimate of the true relative risk (bias away from the null)
- C. The same as the true relative risk



Interpretation of Epidemiologic Studies

In the same case-control study, it was observed that mothers of low birth weight newborns tended to be younger than the mothers of the normal weight newborns. Moreover, smoking rates are known to be higher in younger women in this population.

This scenario would result in which of the following:

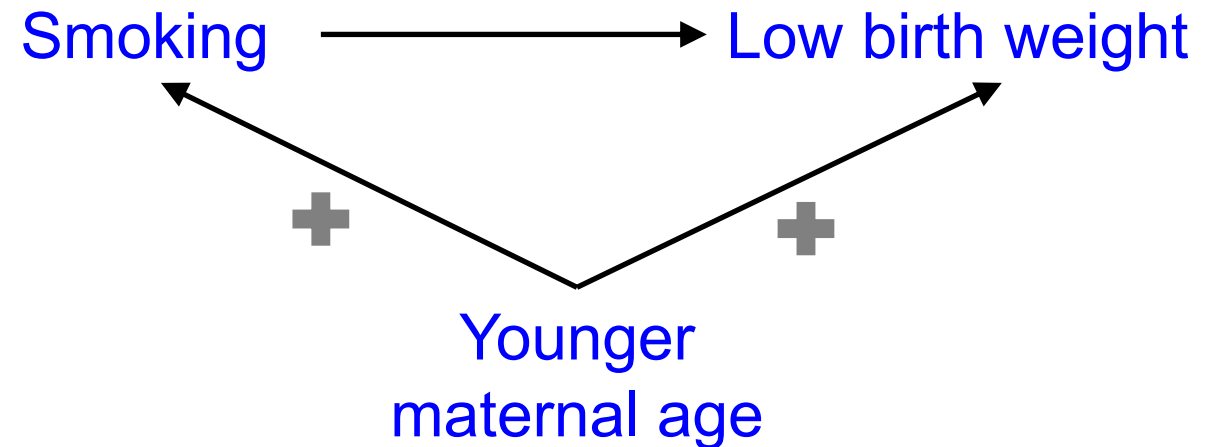
- A. Recall bias and we will underestimate the true relative risk**
- B. Recall bias and we will overestimate the true relative risk**
- C. Confounding and we will underestimate the true relative risk**
- D. Confounding and we will overestimate the true relative risk**



Interpretation of Epidemiologic Studies

In the same case-control study, it was observed that mothers of low birth weight newborns tended to be younger than the mothers of the normal weight newborns. Moreover, smoking rates are known to be higher in younger women in this population.

This scenario would result in which of the following:
D. Confounding and we will overestimate the true relative risk



Types of error

- Chance – always an explanation for our data because we want to draw a conclusion (draw an inference) about the entire population based on a sample
- Overriding principle: the size of the sample on which we are basing conclusions will play a major role in the likelihood that chance is an explanation for our findings



Vignette

In comparing the association between two groups (OR=1.3), the p-value is found to be 0.20. The correct interpretation of this result, testing at the 0.05 level, is:

- A. The null hypothesis is rejected
- B. The association is statistically significant
- C. The association did not occur by chance
- D. The association is compatible with the null hypothesis
- E. Sampling variation is an unlikely explanation of the association



Types of error

- Set up a null hypothesis (H_0): nothing going on, no difference or no association between risk factor and outcome (i.e. $RR = 1$)
- Evaluate the alternative hypothesis: something is happening (i.e. $RR \neq 1$)
- P-value is the probability that the observed data (or data more extreme) would occur due to the effects of chance alone given that the null is true
 - <0.05 – reject H_0 , association is “statistically significant” at the 0.05 level
 - ≥ 0.05 – fail to reject H_0 , association is compatible with the null hypothesis of no association
- P-value does not mean “due to chance” or that chance is ruled out; relates to the likelihood of chance being an explanation for the findings
- Power = $1 - \text{Type II error}$
 - Probability of rejecting the null hypothesis when the alternative hypothesis is true
 - Conventionally set at 80%



Vignette

In comparing the association between two groups (OR=1.3), the p-value is found to be 0.20. The correct interpretation of this result, testing at the 0.05 level, is:

- A. The null hypothesis is rejected
- B. The association is statistically significant
- C. The association did not occur by chance
- D. The association is compatible with the null hypothesis**
- E. Sampling variation is an unlikely explanation of the association



Types of data and types of analysis

Type of Data	Example
Continuous	Age, weight
Dichotomous (binary)	Vital status (dead/alive)
Nominal or categorical	Eye color, blood type
Ordinal	Modified Rankin Scale

Outcome	Explanatory	Analysis (Independent Samples)
Continuous	Dichotomous	t-test
Continuous	Categorical	ANOVA
Continuous	Continuous	Correlation
Dichotomous	Dichotomous	Chi-squared test



MOC REFLECTIVE STATEMENT (BRIEF TAKE HOME NOTES FOR REFERENCE)

Screening

- Sensitivity = $\text{Prob}(T+ | Dz+)$
- Specificity = $\text{Prob}(T- | Dz-)$
- Positive predictive value = $\text{Prob}(Dz+ | T+)$
- Negative predictive value = $\text{Prob}(Dz- | T-)$
- Sensitivity and specificity are influenced (in opposite directions) by the criterion of positivity
- Lead time bias - time by which diagnosis is advanced by screen



MOC REFLECTIVE STATEMENT (BRIEF TAKE HOME NOTES FOR REFERENCE)

Disease Frequency

- Prevalence = proportion existing with disease at a point in time
- Incidence = new cases developing over time among those initially free of disease
- Prevalence $\sim$ Incidence * Duration

Measures of Association

- Relative risk (cohorts/trials) = I_e/I_o ; strength of association
 - Odds ratio (case-control) = ad/bc ; strength of association
- Attributable risk = $I_e - I_o$; public health impact



MOC REFLECTIVE STATEMENT (BRIEF TAKE HOME NOTES FOR REFERENCE)

Types of Studies

- Observational
 - Cohort – enrollment/classified based on exposure; follow for outcomes
 - Case-control – enrollment based on outcome; assess prior exposure
- Randomized clinical trials – exposure allocated by investigator; greater control of confounding; ethical issues

Interpretation of Studies

- Bias: any source of systematic error in the determination of the exposure-outcome association
- Confounding: mixture of effects between the association under study and a third variable, associated with exposure and a risk factor for the outcome
- Chance: sampling variability; assessed through the p-value (probability that the observed data (or data more extreme) would occur due to the effects of chance alone given that the null is true)



REFERENCES

Essentials of Epidemiology in Public Health, 4th Edition
Ann Aschengrau and George R. Seage III
Jones and Bartlett Learning, 2018

Epidemiology in Medicine
Charles H. Hennekens and Julie E. Buring
Little, Brown and Company, 1987

Epidemiology: An Introduction, 2nd edition
Kenneth J. Rothman
Oxford University Press, 2012

