Epidemiology
Epidemiology

Bias

A systematic error in a study which causes an incorrect estimate of the association between exposure and disease.

Selection bias:
Error that occurs when the groups of the study differ systematically.

Randomization is used to limit selection bias.

Observation bias:
Error that occurs during data collection, measurement and analysis due to prior knowledge or subjective feelings of the observer.

Blinding is used to limit observation bias.
INTENTION-TO-TREAT ANALYSIS

In a prospective study, patients are analyzed in the group that they were randomly assigned, regardless of whether or not they received the intended intervention.
INTENTION-TO-TREAT ANALYSIS

OH WHOOPS! WAS I SUPPOSED TO TAKE THAT PILL?

IN AN INTENTION-TO-TREAT ANALYSIS, THIS GUY WILL BE INCLUDED IN THE DATA ANALYSIS FOR THE INTERVENTION GROUP, EVEN THOUGH HE NEVER ACTUALLY RECEIVED IT.
NUMBER NEEDED TO TREAT (NNT)

The number of patients that need to be treated in order for one of them to benefit who otherwise would not have benefitted without the treatment

NNT = 4

YAY! I'M HEALED!
Reliability & Validity

- NOT RELIABLE
- NOT VALID
- RELIABLE
- NOT VALID
- NOT RELIABLE
- VALID

INTERNAL VALIDITY, BUT NO EXTERNAL VALIDITY

INTERNAL VALIDITY AND EXTERNAL VALIDITY
SENSITIVITY AND SPECIFICITY
D-DIMER FOR PULMONARY EMBOLISM
A SENSITIVE TEST, BUT NOT SPECIFIC
TRUE POSITIVE
FALSE POSITIVE
FALSE NEGATIVE
TRUE NEGATIVE
PATIENTS WITH A PE
PATIENTS WITHOUT A PE
PATIENTS WITH A POSITIVE D-DIMER

NITRITES IN URINE FOR UTI
A SPECIFIC TEST, BUT NOT SENSITIVE
TRUE POSITIVE
FALSE POSITIVE
FALSE NEGATIVE
TRUE NEGATIVE
PATIENTS WITH NITRITES IN URINE
PATIENTS WITH A UTI
PATIENTS WITHOUT A UTI
**The Truth**

| Test Result | Positive | Negative |
|-------------|----------|----------|
| Positive    | True Positive (TP) | False Positive (FP) |
| Negative    | False Negative (FN) | True Negative (TN) |

**Positive Predictive Value (PPV)**

\[
\text{PPV} = \frac{A}{A + B}
\]

**Negative Predictive Value (NPV)**

\[
\text{NPV} = \frac{D}{D + C}
\]

**Sensitivity**

\[
\text{Sensitivity} = \frac{A}{A + C}
\]

**Specificity**

\[
\text{Specificity} = \frac{D}{D + B}
\]

**Likelihood Ratio (Fagan) Nomogram**

**Positive Likelihood Ratio (+LR)**

\[
\text{+LR} = \frac{\text{Sensitivity}}{1 - \text{Specificity}}
\]

**Negative Likelihood Ratio (-LR)**

\[
\text{-LR} = \frac{1 - \text{Sensitivity}}{\text{Specificity}}
\]

**Pre-Test Probability**

**LR**

**Post-Test Probability**
When exposed to pollution, there is a 75% (3/4) risk of contracting a disease.

When exposed to pollution, there are 3 to 1 odds of contracting a disease.

**Relative Risk**
\[
\frac{A}{A+B} \div \frac{C}{C+D}
\]
A ratio of two probabilities.

**Odds Ratio**
\[
\frac{A/B}{C/D} = \frac{AD}{CB}
\]
A ratio of two odds (as the name suggests).
TYPE I & TYPE II ERROR

TRUE POSITIVE
THIS IS A HORSE
(OF COURSE)

FALSE POSITIVE
THIS IS A HORSE
(TYPE I ERROR)

FALSE NEGATIVE
THIS IS NOT A HORSE
(TYPE II ERROR)

TRUE NEGATIVE
THIS IS NOT A HORSE
(OF COURSE)
The Truth

Horse 1-β
Not a Horse α

Decision

Horse β
Not a Horse

Type I Error α
Type II Error β
Power 1-β
HIERARCHY OF EVIDENCE

**CASE REPORT**
GRADE: C+
A detailed review of the symptoms, signs, diagnosis, treatment and follow-up of an individual patient

**CASE SERIES**
GRADE: B-
A group of case reports where patients received similar treatment

**CROSS-SECTIONAL STUDY**
GRADE: B
An observational study that analyzes population data at a given point in time
HIERARCHY OF EVIDENCE

CASE-CONTROL STUDY
GRADE: B+

Patients with a disease (cases) are matched with patients without the disease (controls) and data is reviewed retrospectively to determine if there is a common exposure.

Exposed

(+)-Disease

Unexposed

(-)-Disease
Hierarchy of Evidence

Cohort Study

Grade: A–

A longitudinal observational study where exposed and unexposed patients are followed prospectively to determine if there is any difference in outcome.
HIERARCHY OF EVIDENCE

RANDOMIZED CONTROL TRIAL

GRADE: A

A study where patients are allocated at random to receive a particular intervention and followed to determine if there is any difference in outcome.
Systematic Review
Grade: A+

An extensive review of multiple study results which provides a high level of evidence for a particular intervention.
META-ANALYSIS
GRADE: A++

COMBINES DATA FROM MULTIPLE RANDOMIZED CONTROL TRIALS AND USES STATISTICS TO SUPPORT A PARTICULAR INTERVENTION
20.1 Epidemiology Review Questions

1. What best describes the median of a dataset?
   A. The sum of all values in a dataset, divided by the number of values
   B. The middle value when a dataset is ordered from least to greatest
   C. The sum of the most frequent middle value, divided by the number of middle values
   D. The number that occurs most frequently in the dataset

2. You are performing a small cross-sectional study to assess if there is an association between use of cannabidiol (CBD) oil and hypertension. The age distribution of patients is 70, 52, 76, 68, 64, 51, 52, 63, 77, 68, and 52. What is the mode?
   A. 52
   B. 63
   C. 64
   D. 68

3. You are designing a study to assess the effectiveness of a new medication for benign paroxysmal positional vertigo (BPPV). What is the best definition of the number needed to treat (NNT)?
   A. The number of patients one needs to treat with an intervention to achieve a beneficial patient outcome
   B. The number of patients one needs to treat with an intervention to achieve a high patient satisfaction score
   C. The number of patients one needs to treat with an intervention to have a non-beneficial outcome
   D. The number of patients needed in a study to achieve external validity

4. A researcher is performing a study of a new government-funded artificial intelligence laser-guided diagnostic test for freckles. He examines his results which includes pictures of the skin of the study subjects. He finds that a lot of tattoos were classified as freckles by the machine. How would one classify those test results?
   A. True positives
   B. False positives
   C. False negatives
   D. True negatives
5. Researchers are designing a new rapid nasal swab test to diagnose a virus. The test is almost always correct (the patient has the virus when the result is checked by the gold standard test) when it results positive, but misses most patients who have the virus. How would one best describe the test’s validity?

A. High sensitivity, high specificity
B. Low sensitivity, low specificity
C. High sensitivity, low specificity
D. Low sensitivity, high specificity

6. Which of the following is the best definition of a pretest probability?

A. The probability of disease in the local community based on your personal patient experience
B. The probability of disease worldwide based on your personal patient experience
C. The probability of disease existing in a patient after considering testing results, patient history, and clinical exam data
D. The probability of disease existing in a patient after considering disease prevalence, patient history, and clinical exam data

7. You are conducting a case-control study to identify the cause of an outbreak of ciguatera fish poisoning in a small fishing town with 30 case-patients and 40 controls. Among the 30 patients with ciguatera, 29 reported eating at the local fish taco stand, while only 2 of the controls reported eating there. What is the odds ratio for the association between eating at the fish taco stand and contracting ciguatera?

A. 0.95
B. 0.97
C. 36.5
D. 551

8. Which of the following is an example of a type II error?

A. A patient thinks he needs antibiotics for his seasonal allergies.
B. A patient is told he or she has cancer when the CT scan finding is artifact.
C. A patient is told he or she does not have coronavirus when he or she actually has it.
D. The laboratory mixes up the stool samples between patients.

9. Which of the following options is considered the highest level of evidence?

A. A randomized single center trial
B. A cohort study
C. A randomized multicenter controlled trial
D. A systematic review

10. Two groups of test subjects are followed to see how many develop lung cancer. One of these groups vapes while the other does not. Which of the following best describes this study design?

A. Case series  
B. Cohort study  
C. Cross-sectional study  
D. Case-control study

20.2 Epidemiology Review Answers

1. B  
2. A  
3. A  
4. B  
5. D  
6. D  
7. D  
8. C  
9. D  
10. B