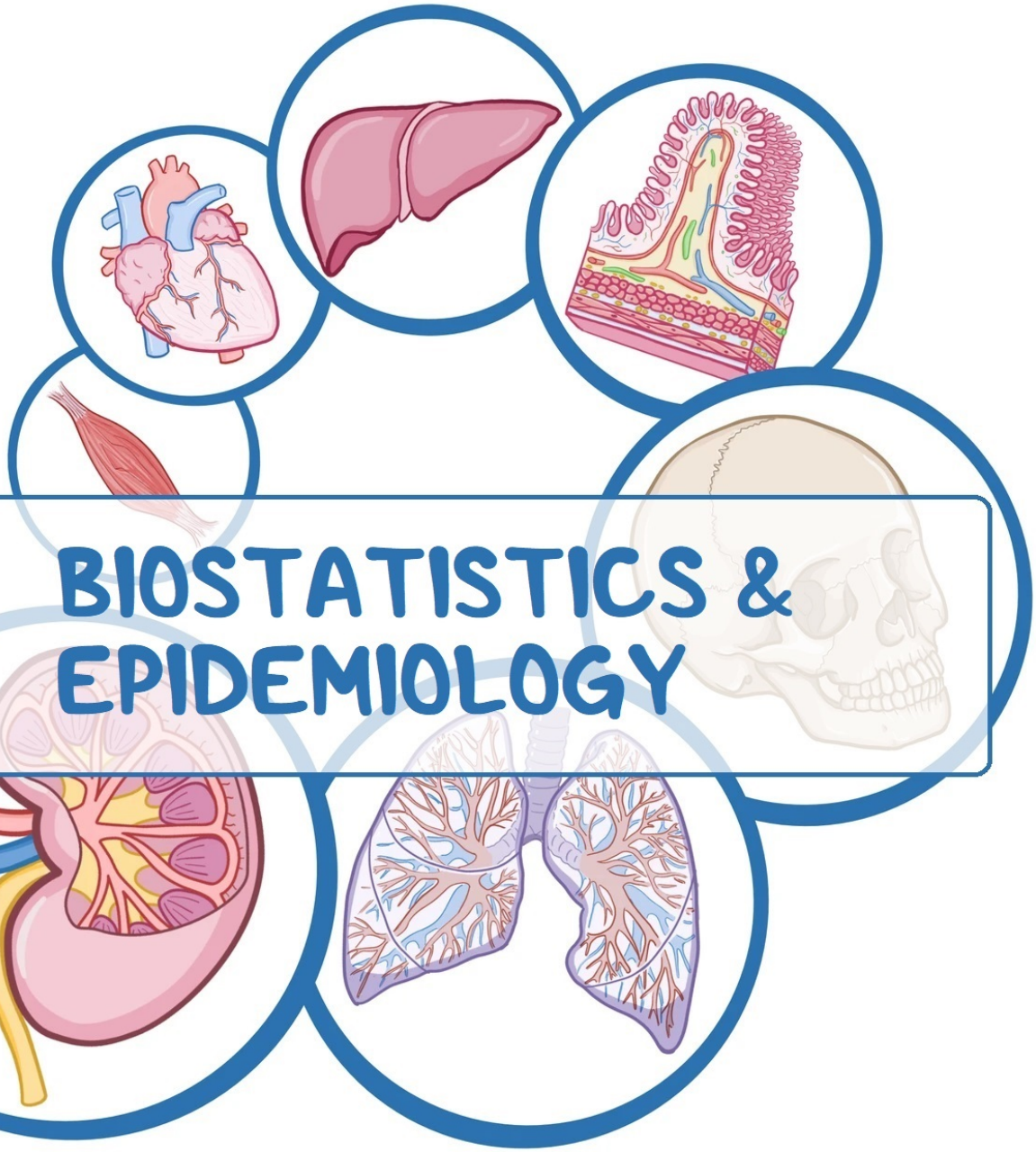
PHYSIOLOGY



BIOSTATISTICS & EPIDEMIOLOGY

HIGH-YIELD
NOTES

AfraTafreeh.com

Table of Contents

NOTES
INTRODUCTORY BIOSTATISTICS

INTRODUCTION TO BIOSTATISTICS

osms.it/intro-biostatistics

- *Statistics:* process of collecting, organizing, analyzing data set variables
- *Biostatistics:* focus on data related to living things
- *Descriptive statistics:* summarizes, describes population information
- *Inferential statistics:* examines relationships between two/more variables → applies results of sample population to target population

POPULATION & SAMPLE

Population

- Group (people, specimens, events) with defined criteria (e.g. October–March emergency room visits)
- *Parameter:* numerical population description (e.g. range, mean, standard deviation)
  - μ = population mean
  - σ = population standard deviation

Sample

- Subset drawn from population (e.g. influenza-related October–March emergency room visits)
- Represents population → inferences can be made about population
- *Statistic:* numerical sample description (e.g. range, mean, standard deviation)
- $\overline{X}$ = sample mean
- SD = sample standard deviation
- *Sampling error:* sample does not accurately reflect population
  - Usually due to wide variation within sample
  - ↑ sample size helps avoid sampling error
- *Selection bias:* sample does not accurately reflect population
  - Occurs when precautions to obtain representative sample are not used
  - Randomization helps eliminate bias

Case (data point)

- Single observation (e.g. one individual visiting emergency room for influenza symptoms)

TYPES OF HYPOTHESES

Null hypothesis (H_0)

- States that there is no relationship between variables
- Any observed relationship due to chance (e.g. no relationship between body mass index (BMI), hypertension)

Alternative hypothesis (research hypothesis)

- States expected relationship between variables (e.g. relationship between BMI, hypertension)

Hypothesis testing

- Statistical methods used to determine relationship strength between variables, how much of observed relationship is due to chance, significance of observations
- *Statistical significance:* relationship between variables is caused by something other than chance
- Usually defined by a p-value of < 0.05 (5%); "*p*" stands for "probability"
  - *Type 1 error:* probability of incorrectly rejecting null hypothesis (i.e. concluding significant relationship between

variables when there is not)
  - *Type 2 error:* incorrectly accepting null hypothesis (i.e. concluding there is no significant relationship between variables, missing present association)
- *Clinical significance:* practical importance of study results that may not be statistically significant

RELIABILITY & VALIDITY

- Measurement characteristics used to collect data

Validity: accuracy

- Instrument actually measures variable (concept, construct) it is supposed to measure (e.g. urine dipstick accurately detects proteinuria)
- Valid instrument must be reliable

Reliability: repeatability

- Instrument consistently yields same results with repeated measurements (e.g. urine dipstick reliably detects proteinuria with each measurement)
- Reliable instrument may/may not be valid

TYPES OF VARIABLES

- *Variable:* defined characteristic being studied; can assume different values
- *Independent variable:* manipulated (treatment) variable
- *Dependent variable:* outcome variable; influenced by independent variable
  - What is effect of X (independent variable) on Y (dependent variable); how is X related to Y?
  - E.g. what is the effect of lipid-lowering drug (X) on individual's cholesterol level (Y)?

GRAPHIC DESCRIPTION OF DATA

- When values are plotted on graph → variety of frequency distributions (curves) result
- *Properties of distributions:* central tendency, dispersion

Normal (Gaussian) curve

- Symmetrical distribution of scores around mean
  - Forms classic bell shape
  - Values lie within two standard deviations of mean
  - Most natural phenomena show this type of distribution
  - Parametric tests utilized in research

Non-Gaussian curve

- Asymmetrical distribution of scores around mean
  - Skewed (negatively/positively) curve
  - Kurtotic (flat/peaked) curve (leptokurtic—thin, positive kurtosis; platykurtic—flat negative kurtosis)
  - Nonparametric tests utilized in research

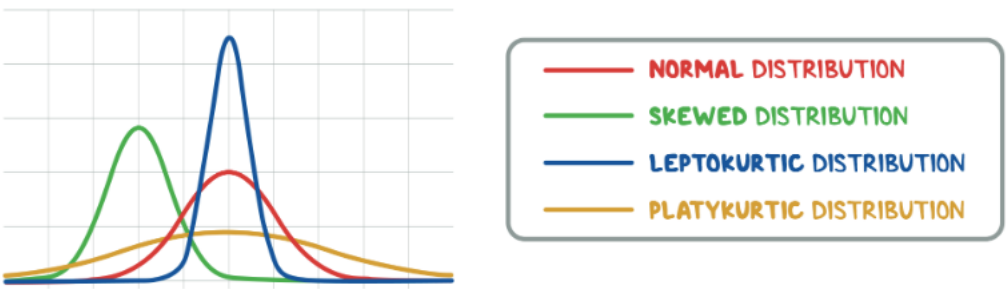


Figure 5.1 Visualization of normal (red), skewed (green) and kurtotic (blue and yellow) distributions.

MEAN, MEDIAN, MODE

osms.it/mean-median-mode

- Central tendency measures
- More curve symmetry → more alike mean, median, mode

Mean ($\overline{X}$)

- Central value calculated by adding each value in data set → dividing by total number of data points
- *Expressed as formula:* total sum of individual data points X_1, X_2, ..., X_n, divided by n (number of data points)

$$\overline{X} = \frac{(X_1 + X_2, ..., + X_n)}{n}$$

$$\frac{17+19+20+20+61+61+62}{7} = \frac{260}{7} = 37.14$$

- Can be influenced by an extreme value (outlier) → skewed data

Median

- Calculates central value when possible outliers present
- Divides set of data into two halves
  - Half of values > median, half < median
- Most commonly used expression of central tendency
- Arrange data in order of magnitude → find midpoint

17 19 20 20 61 61 62 100

- Odd number of values → one "middle" number
- Even number of values → two middle-values values (20, 61)
  - Calculate median by averaging two values: (20+61)/2 = 40.5

Mode

- Central value appearing most often in data sequence
  - Bimodal (two modal), trimodal (three modes), amodal

17 19 20 20 61 61 62 100

  - Bimodal dataset with two mode values of 20, 61
- Not affected by outliers

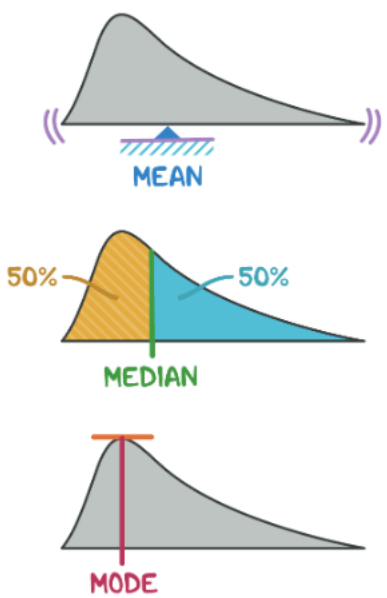


Figure 5.2 Mean, median, and mode in a skewed curve.

PROBABILITY

osms.it/probability

- Relative likelihood that event will/will not occur
- To calculate chance that event/outcome will occur → divide number of times event happened by number of times event could have happened
  - E.g. event A is rolling a die and getting a three
  - Since a die has six sides, there are six possible numbers, so the probability (P) of rolling a three is 1/ 6, or 0.167 (16.7%)

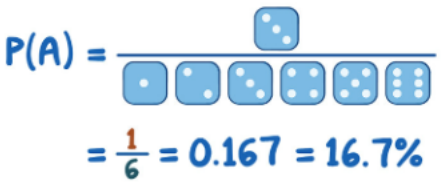


Figure 5.3 Probability of rolling a three on a six-sided die.

RULES

Rule 1

- Probability of event A can range anywhere from 0% to 100%
  - $0 \leq P(A) \leq 1$

Rule 2

- Sum of probabilities of all possible outcomes = 1

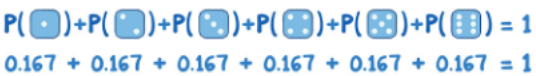


Figure 5.4 Visualization of Rule 2.

Rule 3 (complement rule)

- Probability that event will not occur = 1 minus probability that it does occur
  - P = 1 – P(A)

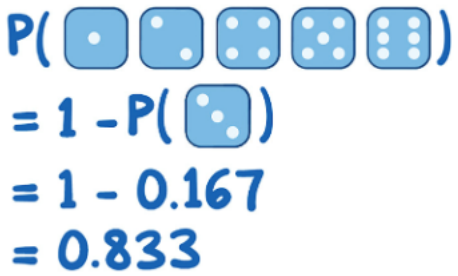


Figure 5.5 Probability of not rolling a three = 1 - P(rolling a three).

Rule 4

- Probability of two disjoint (mutually exclusive) events = the sum of the first event plus the second event
  - P(A or B) = P(A) + P(B)

Rule 5

- Probability for two not disjoint (not mutually exclusive) events = sum of the probability of event A and the probability of event B, minus the probability of event A and B together
  - P(A or B) = P(A) + P(B) – P(A and B)

Rule 6

- Probability of two independent events = probability of the first event multiplied by the probability of the second event
  - P(A and B) = P(A) x P(B)

Rule 7

- Conditional probability (probability of event A, given what happens in event B) = probability of event A and event B divided by probability of event B

Rule 8

- Probability of events A, B = probability of event A multiplied by conditional probability of event B given event A occurred

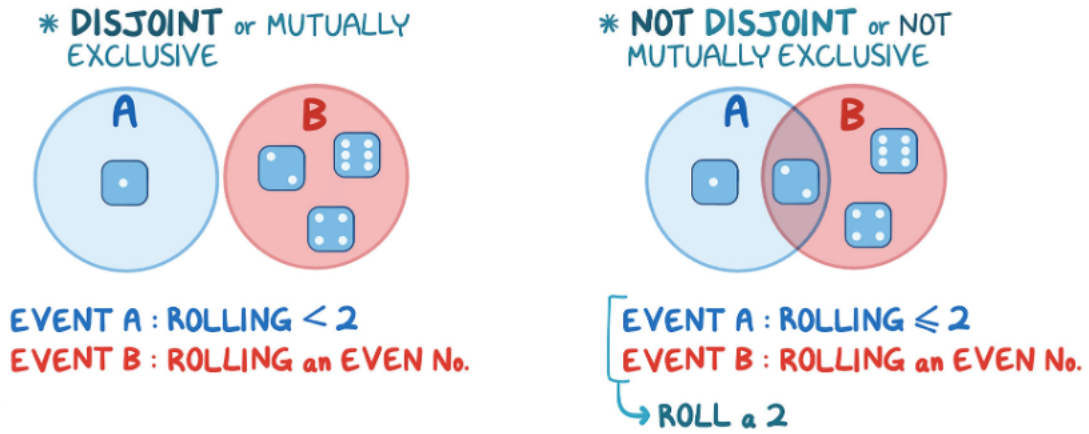


Figure 5.6 A visualization of the difference between mutually exclusive and not mutually exclusive events.

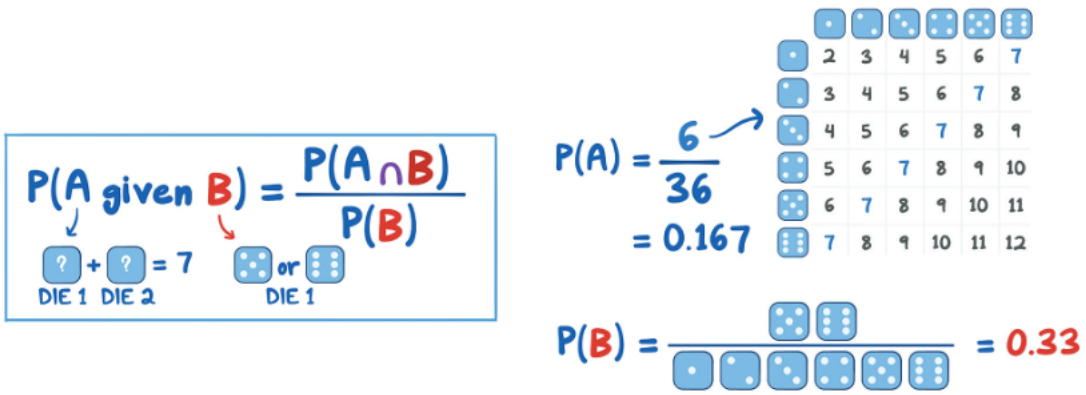


Figure 5.7 Rule 7, conditional probability: determining P(A) and P(B) when event A depends on event B. In this case, we are finding the probability that the roll of two dice adds up to seven (event A) given that the first die is either a five or a six (event B). Once P(A) and P(B) are known, they are used to solve for P(A given B).

RANGE, VARIANCE, & STANDARD DEVIATION

osms.it/range-variance-standard-deviation

- Measures distribution of variables

Range

- Difference between highest, lowest value
- E.g. Range of individuals' cholesterol levels
  - 130, 150, 152, 158, 165, 289, 354
  - Range 354 - 130 = 224mg/dL
- E.g. individual weight (in kg)
  - 10 + 45 + 50 + 55 + 90
  - Range = 90 - 10 = 80

Variance

- Sum of squared deviations from mean, divided by number of distributions

$$\sigma^2 = \frac{\Sigma(x-x)^2}{n}$$

- E.g. variance of individual weight (in kg)
  - $(10 - 50)^2 + (45 - 50)^2 + (50 - 50)^2 + (55 - 50)^2 + (90 - 50)^2 / (5) = 650\ kg^2$

Standard deviation (SD)

- Square root of variance

$$\sigma = \sqrt{\frac{\Sigma(x-x)^2}{n}}$$

  - E.g. SD of individual weight: $\sqrt{650} = 25.5kg$
- In Gaussian curve
  - *68 - 95 - 99 rule*: 68% of data points lie within 1 SD from mean; 95% lie within 2 SD, 99% lie within 3 SD
- Z-score = number of SD data point is away from mean
  - Data point minus the population mean, divided by the population standard deviation

$$\frac{x-\mu}{\sigma}$$

  - E.g. blood glucose population mean = 90g/dL, SD = 20g/dL, data point = 130g/dL (130 - 90 / 20 = 2)
- Coefficient of variation (CV) = SD/mean; also expressed as percentage, obtained by multiplying the CV by 100

TYPES OF DATA

osms.it/types-of-data

- Determining type of data to be collected helps establish which sort of distributions can logically be used to describe variable

Nominal data

- Can assume one of a limited number of possible values (e.g. ABO blood types)
  - No meaningful rank order; no median, mean, standard deviation; mode used for analysis
  - Includes dichotomous variables (e.g. normal, abnormal)

Ordinal data

- Ordered in meaningful way (e.g. systolic murmur ranking from 1–6)
  - Follows order, but quantitative differences not clear (do not indicate degree of difference between observations)
  - Median, mode can be used; mean usually not suitable to describe sample/population

Discrete data

- Measured in whole numbers (no decimal values)
  - E.g. number of pregnancies

Continuous data

- Can take on infinite number of value (e.g. weight, height, blood glucose)
  - Mean, median, mode, standard deviation can be calculated

Interval data

- Indicates meaningful quantitative difference between two values; values can be placed in clear, logical order
  - E.g. temperature on Celsius/Fahrenheit scale; difference between 90° and 60° measured as 30°
  - Arbitrary zero point
  - Mean, median, mode, standard deviation can be calculated

Ratio data

- Has absolute, meaningful zero point
- Can use multiplication, addition, subtraction to calculate ratios
- Mean, median, mode using ratio data

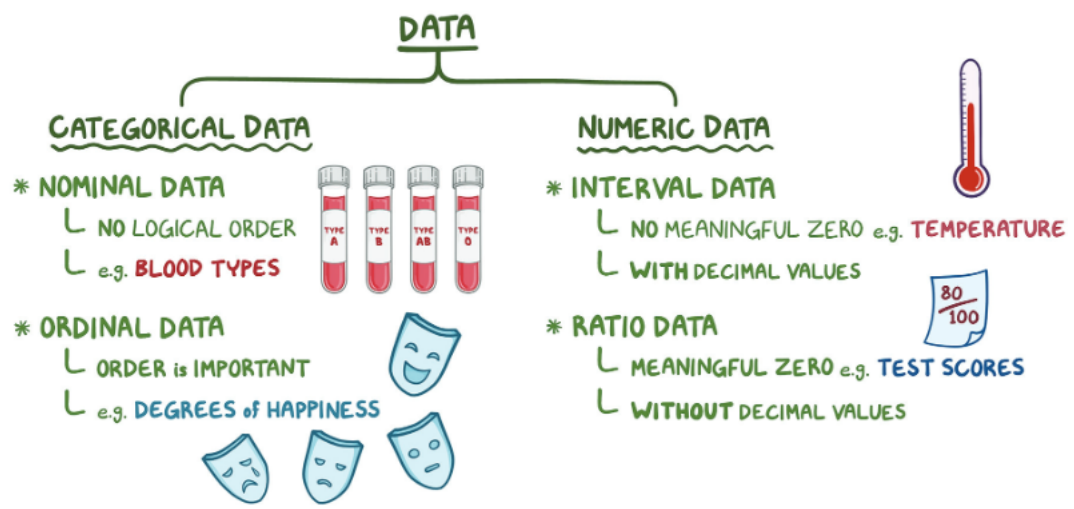


Figure 5.8 Types of data.

NOTES
CAUSATION & VALIDITY

CAUSALITY

osms.it/causality

- Consequential relationship between two events (e.g. A caused B)
  - *Contrast with correlation:* association between two events
- Consequential relationship may be direct/indirect
  - *Direct:* event caused direct consequence which → effect (A → B)
  - *Indirect:* initial event → another event → final effect (A → x → y → B)
- Correlation is not equal to causation
  - Two correlated events may seem to have consequential relationship; sometimes due to random chance/external factors/confounding (noncausal) variables

ESTABLISHING CAUSALITY

- To establish causality between set of events, relationship must meet following criteria

Temporality

- Cause happened before effect
  - Event A followed by Event B
  - *Example:* smoking → lung cancer

Strength of association

- Relational closeness between two events
  - Measured by relative risk, odds ratio, correlations, etc.
  - *Example:* how closely is smoking related to developing lung cancer?

Dose-response relationship

- More exposure to cause → greater effect
  - Longer exposure to Event A → more risk of Event B
  - *Example:* the longer you smoke, the higher your risk of developing lung cancer

Biologic coherence

- Causal mechanism for effect agrees with current knowledge
  - *Example:* factually known that cigarettes contain carcinogenic agents

Biologic plausibility

- Proposed mechanism of effect makes sense according to current knowledge
  - *Example:* because we know cigarettes contain carcinogenic agents, it makes sense that cigarette-smoke exposure → higher probability of developing lung cancer

Consistency with other knowledge

- Association has been shown repeatedly
  - *Example:* it has been repeatedly proven that smoking confers higher risk of developing lung cancer

Specificity

- Chances that effect is due to other causes
  - *Example:* can there be another explanation for developing lung cancer besides exposure to cigarette smoke?

Experimental evidence

- When you remove cause, effect disappears
  - *Example:* if you stop smoking, your risk of developing lung cancer decreases

Analogy

- Similar events have been proven to cause similar effects
  - *Example:* smoking other substances has

been known to cause lung pathology

CAUSAL RELATIONSHIP TYPES

Necessary and sufficient

- Presence of A required, present in adequate amounts to cause B
  - *Example:* autosomal dominant mutation with complete penetrance both necessary, sufficient for disease to develop

Necessary but not sufficient

- Presence of A required, not present in adequate amounts to cause B
  - *Example:* heat required to cause burn, however, low heat will not cause burn; it is necessary but not sufficient

Not necessary but sufficient

- Presence of A not required, but is enough to cause B
  - *Example:* gunshot to head sufficient to cause death, however, not necessary, as there are many other causes of death

Not necessary and not sufficient

- Presence of A not needed nor enough to cause B
  - *Example:* urinary infection not necessary nor sufficient to cause pelvic inflammatory disease; urinary infection can be present without pelvic inflammatory disease, individual can have pelvic inflammatory disease without having urinary tract infection

BIAS

osms.it/bias

- Error in one step of study design/ conduction/analysis → results interpretation that is different from truth
  - Many types of biases, no common classification

SELECTION BIAS

- Errors made when choosing/following population to be studied
  - Can occur at different stages of study
  - Most commonly occurs when chosen sample is not representative of population

MEASUREMENT BIAS

- AKA information bias
- Errors made when measuring data/results of interest
  - Most commonly results in results misclassification which can be differential/non-differential

Differential misclassification

- Error in measurement more likely to occur in one group than another
  - Results of one group will be inherently different to other group's results
  - *Example:* blood glucose levels of groups measured by different machines; one gave accurate results, other reported inaccurate results

Non-differential misclassification

- Measurement error likely to have occurred in both groups
  - Results among two groups will not differ greatly
  - *Example:* machine used to determine blood glucose levels for both groups was inaccurate

OTHER BIAS TYPES

- Information gathering, management can → other bias types
- AKA information bias

Procedure bias

- People allocated to different groups not treated identically
  - Usually due to lack of blinding

- *Example:* people in one group spend more time in hospital than other group

Recall bias

- Awareness of event/effect influences individual's recall of cause
  - Most common in retrospective studies
  - *Example:* after a person with cancer knows that radiation exposure is a cancer development risk factor, the person may place more emphasis on exposure to radiation than someone without cancer

Lead-time bias

- Early diagnosis extends follow-up period, making it seem as if event being studied took longer to progress
  - *Example:* early cervical cancer detection may make it seem as if cancer is less aggressive because of more time spent living with diagnosis

Observer-expectancy bias

- When belief in intervention's effectiveness interferes with reported treatment outcome
  - *Example:* researcher's belief in drug efficacy may interfere with reported results

CONFOUNDING

osms.it/confounding

- Occurs when external event is related to possible cause, outcome of interest but is not on causal pathway
- Example: study exploring relationship between exercising, overall health, we know that
  - Exercising known to improve overall health
  - Exercising associated with healthy lifestyle, but is not result of healthy lifestyle

INTERACTION

osms.it/interaction

- Combination of two/more factors changes disease incidence compared to influence they would have had individually
  - Describes way multiple factors interact to produce event

Synergism

- Refers to potentiation effect multiple factors may have on one another
- *Example:* 2 + 2 = 5

Antagonism

- Refers to inhibition effect multiple factors may have on one another
- *Example:* 2 + 2 = 3

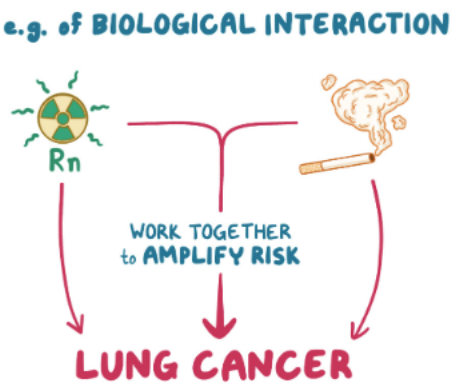


Figure 6.1 Biological interaction is when two exposures, like radon gas and cigarette toxins, work together to influence an outcome, like lung cancer.

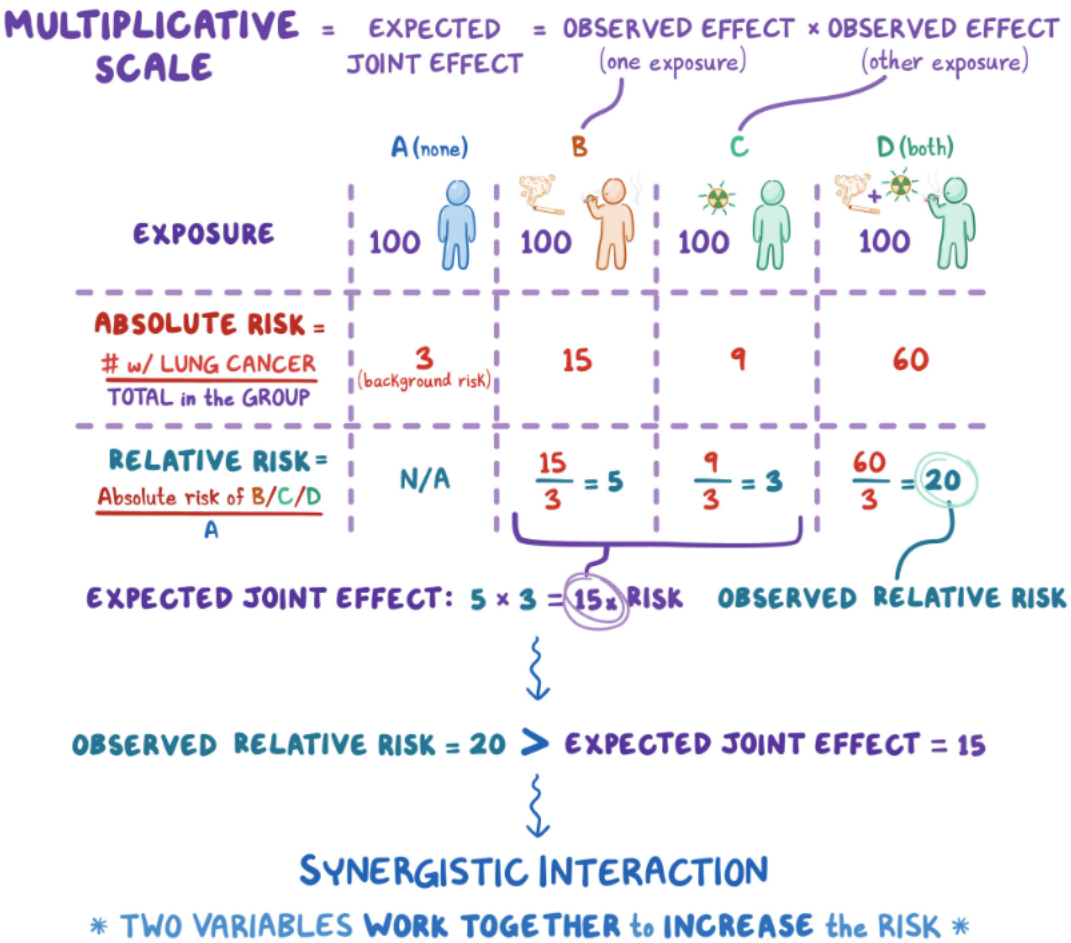


Figure 6.2 A graph representing data collected from four groups with 100 people per group: those with no exposure to radon or cigarette toxins (A), those with exposure to only cigarette toxins (B), those with exposure to only radon (C), and those with exposure to both radon and cigarette toxins (D). The multiplicative scale was used to calculate the expected joint effect of radon and cigarette toxins based on their independent effects (columns B and C). These two exposures are said to have a synergistic interaction because observed relative risk > expected joint effect. If observed relative risk had been < expected joint effect, the interaction would have been antagonistic.

NOTES
COMMUNITY HEALTH

DYNAMICS OF OUTBREAKS

- Outbreak: sudden increase in disease occurrence in a specific time, place, population (e.g. outbreaks of foodborne-related norovirus acute gastroenteritis)
- Infective outbreaks depend on causative pathogen characteristics (such as mode of transmission)

MODES OF INFECTIOUS DISEASE TRANSMISSION

osms.it/transmission

TRANSMISSION

- The passing of a pathogen-causing communicable disease from an infected host to another individual/group

MODES OF TRANSMISSION

- Depends on responsible organism's characteristics

Direct

- Interpersonal contact → infected individuals spread disease
  - One-to-one (e.g. venereal infection with sexual intercourse)
  - One-to-multiple (e.g. influenza with a violent sneeze in a crowded environment)

Indirect

- Common vehicle (e.g. contaminated air, water/food supply, needle-sharing)
- Vectors (e.g. mosquito/tick)

OUTBREAK INVESTIGATIONS

osms.it/outbreak-investigations

CHARACTERISTICS OF AN OUTBREAK

- *Explosive:* in epidemic curve, there is a fast, abrupt rise in number of cases, followed by fast, abrupt fall
- *Indirect transmission:* infection limited to individuals who share common exposure
- *Direct transmission:* often impossible to associate new cases to primary case (first symptomatic case occurring in defined setting)

STEPS TO INVESTIGATE AN OUTBREAK

Validate outbreak's existence in a population

- Define number of cases (numerator)
- Define the extent of the population susceptible to disease (denominator)
- Determine whether number of observed cases is more than expected number of cases
- *Calculate the attack rate*: proportion of an initially disease-free population that develops disease
  - Proportion is used because the individuals in the numerator (those who have the disease) are included in the denominator (the total population)

Investigate cases by looking for interactions of time, person, place

- Are there interactions between variables?

Develop hypotheses

- Consider existing knowledge about the disease, findings from current investigation

Test hypotheses

- Analyze data (e.g. case-control study, laboratory tests such as chemical/ immunological fingerprinting)

Recommend measures for disease control, prevention

- E.g. remove infection source, establish environmental controls (interrupt disease transmission), improve sanitation, immunize susceptible individuals

DISEASE SURVEILLANCE

osms.it/disease-surveillance

- Essential public health tool, aimed at predicting, observing, minimizing outbreaks
- Based on systematic collection, analysis, interpretation of epidemiologic data
- Monitored parameters examples
  - Changes in disease incidence/mortality
  - Changes in quantity of risk factors for a disease in environment
  - Completeness of vaccination coverage
  - Prevalence of drug-resistant organisms

MODALITIES OF SURVEILLANCE

Passive

- Using existing data on reportable diseases such as anthrax, cholera, gonorrhea
- Pros
  - Comparatively inexpensive, easy to develop
  - Areas that require urgent intervention are quickly identified by international comparisons
- Cons
  - Surveillance is not the primary responsibility of case-reporting individuals
  - Local outbreaks may be missed

Active

- Implementing surveillance program (e.g. field visits to clinics, hospitals, communities)
- Pros
  - Reporting more accurate; individuals recruited specifically for surveillance program
  - Local outbreaks are more likely to be identified
- Cons
  - More expensive to develop, maintain

DIFFICULTIES

- Obtaining reliable data in low-income countries → underreporting risk
  - Areas may be difficult to reach
  - Communication with central authorities can be challenging
  - Resources such as diagnostic laboratories not always available

VACCINATION & HERD IMMUNITY

osms.it/vaccination-herd-immunity

HERD IMMUNITY BASICS

- *Herd immunity:* phenomenon in which entire population is indirectly protected against disease when critical percentage of members are immune
  - Immunity can be innate/acquired through vaccination/by naturally recovering from infection
  - The higher the proportion of immune people in a population, the less likely the encounter between a susceptible person and an infected one → chain of infection is disrupted

Conditions

- Host is a single species
- Transmission of the organism must be spread by direct contact
- No reservoir outside the human host
- Infections must induce solid immunity

HERD IMMUNITY & COMMUNITY HEALTH

- The critical percentage of immune individuals needed to achieve herd immunity varies according to disease contagiousness (e.g. 94% in measles [highly communicable] → increased number of individuals need to be immune)
- Because of herd immunity, vaccination programs do not necessitate yield 100% immunization rates, yet can achieve highly effective protection by immunizing critical percentage of a population
- Herd immunity is important for public health because individuals who cannot develop immunity or cannot be vaccinated depend on herd immunity (e.g. newborn infants, individuals with immunodeficiency due to HIV/AIDS, cancer, cancer treatments)

NOTES
EPIDEMIOLOGY MEASURES

DIRECT STANDARDIZATION

osms.it/direct-standardization

STANDARDIZATION

- Methods used to compare health event rates of two/more populations (e.g. mortality rates) by standardizing characteristics responsible for inter-population differences
- E.g. remove confounding variables (age) when comparing two groups' crude mortality rate (CMR) to get age-adjusted mortality rate
  - *CMR:* number of people who died in one group, divided by the group population (100,000 or 1,000)

DIRECT STANDARDIZATION

- Compares differences in health events among two/more populations by calculating age-adjusted rate
- Used when event distribution in each age group within population is known
- Process for calculating direct standardization for age-adjusted mortality rate
  - Choose reference (standard) population (e.g. separate population such as a national-level population)
  - Multiply other population of interest's age-specific mortality rates to number of people in each age group of reference population
  - Add up number of expected deaths from all age groups
  - Calculate age-adjusted mortality rate
  - Compare two age-adjusted mortality rates

INDIRECT STANDARDIZATION

osms.it/indirect-standardization

- Used when number of events/mortality rates in each age group within population is not known
- Process for calculating indirect standardization for age-adjusted mortality rate
  - Choose reference population with known mortality rates
  - Multiply other population of interest's age-specific mortality rates to number of people in each age group of reference population
  - Add up number of expected deaths from all age groups
  - Calculate standardized mortality ratio (SMR)

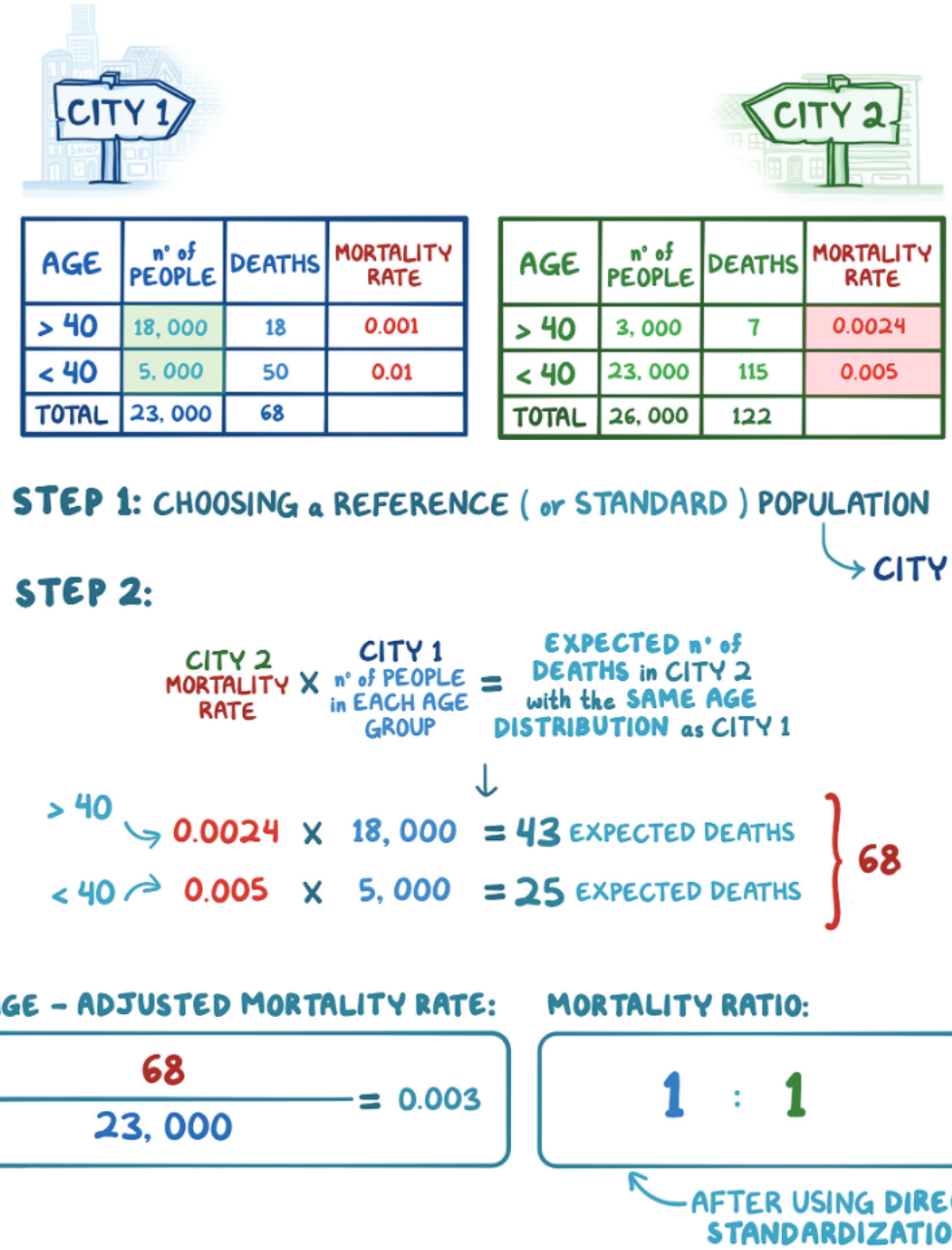


Figure 8.1 Using direct standardization to find the age-adjusted mortality rate for City 2, using City 1 as the reference population.

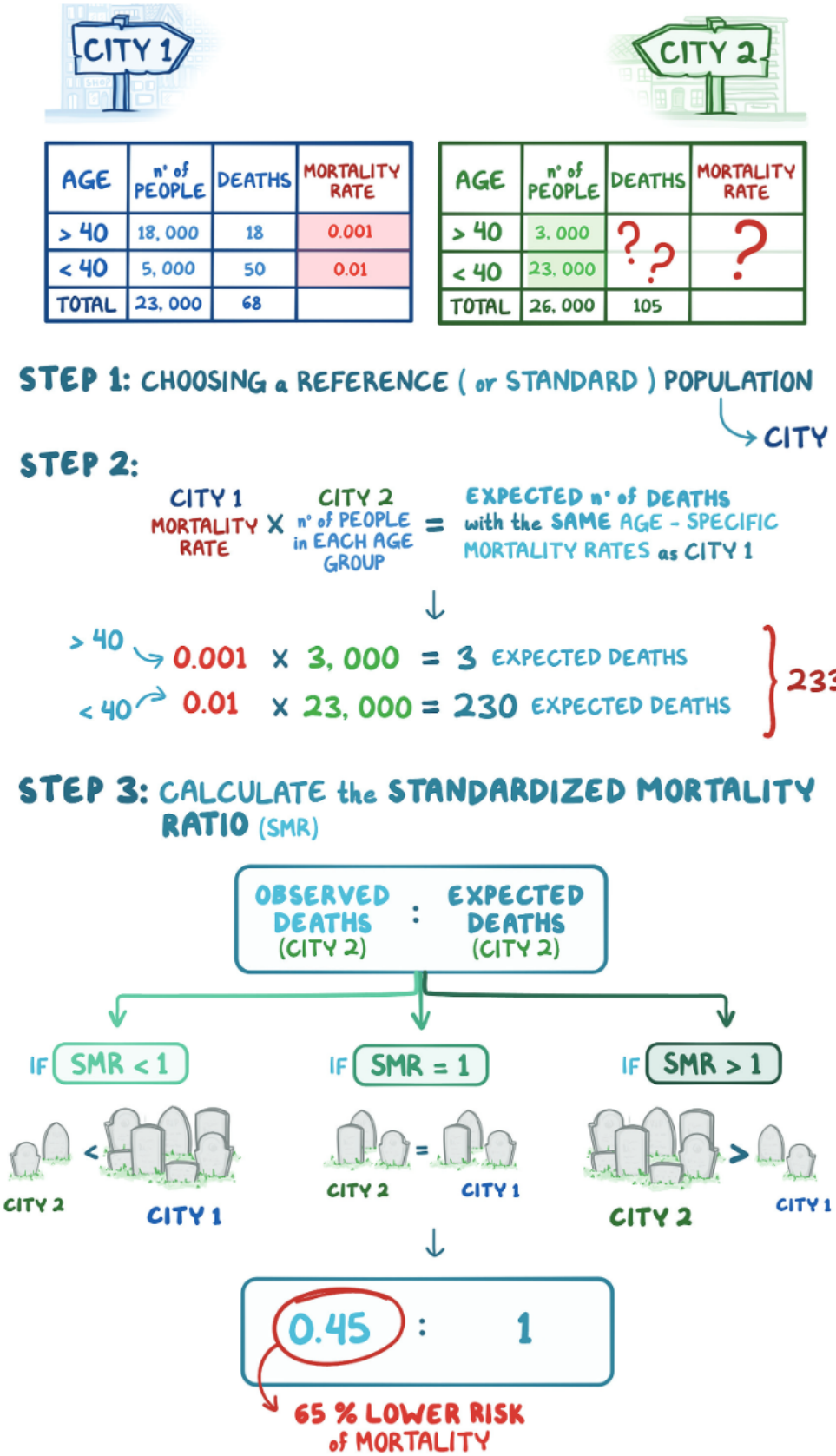


Figure 8.2 Using indirect standardization to find the standardized mortality ratio for City 2, using City 1 as the reference population.

INCIDENCE & PREVALENCE

osms.it/incidence-prevalence

- Measures number of people who have disease
- Reported as population percentage/ratio (e.g cases per 1000)

Incidence

- Number of new disease cases in population over time period (usually one year)
  - Affected by preventive measures (vaccination, diagnostic techniques)

Prevalence

- Number of total (old, new) disease cases in population in particular time point (point prevalence)
  - Shows disease commonness in group of people
  - Affected by cure rate, survival rate, death rate, recurrence

Relationship between incidence and prevalence

- New disease cases (incidence) added to amount of disease present in population (baseline prevalence) → ↑ prevalence
- ↑ death rate, cure rate → ↓ prevalence (↓ total disease cases)
- If incidence > death/cure rate → net ↑ prevalence; if incidence < death/cure rate → net ↓ prevalence

$$\text{prevALence} = \frac{\text{ALL cases}}{\text{population at risk}}$$

$$\text{iNcidence} = \frac{\text{New cases}}{\text{population at risk}}$$

MEASURES OF RISK

osms.it/measures-of-risk

- Probability that event will occur (e.g. disease development risk)

Absolute risk

- Disease incidence in population who have been exposed to specific risk factor
  - E.g. 1 out of 50 (2%) diabetics will develop cardiovascular disease (CVD)

$$\text{Absolute risk} = \frac{\#\text{ of events in a group}}{\#\text{ of individuals in that group}}$$

Relative risk (RR)

- Compares disease development probability between exposed group, unexposed group
  - E.g. smokers' bladder cancer incidence (30%), non-smokers' bladder cancer incidence (3%)
  - RR = 0.3/0.03 = 10
  - Smokers are 10 times more likely to develop bladder cancer

$$\text{Relative risk} = \frac{\text{Probability of event in exposed population}}{\text{Probability of event in unexposed population}}$$

Absolute risk reduction (ARR)

- AKA risk difference
- Outcomes comparison (change in risk)
  - Between population that has received treatment for a disease, population that has not received treatment
- ARR = risk (untreated) - risk (treated)
  - E.g. 4% bladder cancer occurrence in group that receives particular drug, 20% in group that does not receive drug
  - ARR = 0.2 - 0.04 = 0.16 or 16%
  - For every 100 individuals receiving drug, 16 bad outcomes would be avoided

Number needed to treat

- Determines how many individuals should be treated with medication to prevent one person from developing bladder cancer
  - Number needed to treat: 1/0.16 = 6.25
  - About six people should be treated

$$\#\text{ needed to treat} = \frac{1}{\text{Absolute risk reduction}}$$

ODDS RATIO

osms.it/odds-ratio

- Measures association between exposure (e.g. risk factor, health characteristic), outcome (e.g. disease, mortality)
  - E.g. Which group is at higher risk of experiencing an adverse outcome? Does an intervention change risk degree for a group?
- *Used in case-control studies*: case group with identified outcome, control group without identified outcome
- Calculated using 2X2 frequency table
  - Divide odds of disease in exposed individuals by odds of disease in unexposed individuals

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

$$OR = \frac{40/20}{60/80} = \frac{2}{0.75} = 2.66$$

- OR = 1 → exposure does not affect odds of outcome
- OR > 1 → exposure associated with higher odds of outcome
- OR < 1 → exposure associated with lower odds of outcome

2 x 2 FREQUENCY TABLE

	+ (DISEASE)	– (NO DISEASE)
+ (EXPOSED)	**a:** # of exposed individuals (smokers) w/bladder cancer = 40	**b:** # of exposed individuals (smokers) w/o bladder cancer = 60
– (UNEXPOSED)	**c:** # of unexposed cases w/bladder cancer = 20	**d:** # of unexposed cases w/o bladder cancer = 80

ATTRIBUTABLE RISK (AR)

osms.it/attributable-risk

- AKA risk difference/excess risk
- Measures difference in disease risk between exposed population, unexposed population
  - Often used in cohort studies

AR for exposed individuals

$$AR = \frac{A}{A+B} - \frac{C}{C+D}$$

2 x 2 FREQUENCY TABLE

	+ (DISEASE)	– (NO DISEASE)	TOTALS
+ (EXPOSED)	**a:** # of exposed individuals (smokers) w/bladder cancer = 40	**b:** # of exposed individuals (smokers) w/o bladder cancer = 60	**100**
– (UNEXPOSED)	**c:** # of unexposed cases w/bladder cancer = 20	**d:** # of unexposed cases w/o bladder cancer = 80	**100**
	60	**140**	**200**

$$AR = \frac{40}{100} - \frac{20}{100} = 0.4 - 0.2 = \frac{20}{100}$$

$$\frac{AR}{\text{incidence in exposed}} \times 100$$

$$\frac{20}{40} \times 100 = 50\%$$

- 50% of bladder cancer incidence → attributable to smoking in exposed population

AR for population (PAR)

PAR = incidence in population - incidence in unexposed

$$\text{PAR} = \frac{60}{200} - \frac{20}{100} = 0.3 - 0.2 = \frac{10}{100}$$

$$\frac{10}{20} \times 100 = 50\%$$

MORTALITY RATES & CASE-FATALITY

osms.it/mortality_rates_case-fatality

- *Mortality rate:* death incidence in population over period of time

Annual mortality rate

- Mortality rate from all causes (crude death rate)
  - Calculated by taking total number of deaths from all causes in one year divided by total number of people at risk in population at mid-year
- Annual mortality = total number of deaths (850) ÷ total number of people at risk in population at mid-year (500,000) = 0.0017

- *Percent:* 0.0017 x 100 = 0.17%
- *Per 100,000:* 0.0017 x 100,000 = 170 (170 death per 100,000 people during year)

Population-specific mortality rate

- Mortality rate for specific sub-population (e.g. biologically-female individuals; cancer-related deaths, neonatal mortality)
  - E.g. neonatal mortality rate = number of deaths among children < 28 days old (during given time interval) ÷ number of live births (during same time interval) x 1,000

Case-fatality rate

- Percent of people that die within certain period of time post-diagnosis
  - Calculated by dividing number of post-diagnosis deaths by total number of diagnosed individuals, multiplied by 100
- Measures disease severity

$$\text{Case mortality rate from disease A} = \frac{40}{250} = 0.16 = 16\%$$

DALY & QALY

osms.it/DALY-QALY

- *Disease burden measurement:* impact of health problem on individual/population

Disability-Adjusted Life Years (DALY)

- Determines disease burden according to years of life, or to compare specific intervention's effect (e.g. new medication reducing diabetes risk)
  - Morbidity, mortality combined into single metric
- *DALY:* years of lost life due to premature death (YLL) + years lived disability (YLD)
  - *YLL:* number of deaths (N) x standard life expectancy at age of death in years (L)
  - *YLD:* number of incident cases (I) x disability weight (DW) X average duration of disability in years (L)
  - *DW:* reflects disease severity on 0 (perfect health) to 1 (dead) scale

Quality-Adjusted Life Years (QALY)

- Determines disease burden according to quality of years of life, relative value of interventions (e.g. cost-utility analysis); guides healthcare-resource prioritization
- Measures years of life with illness/disability (considered less than year of healthy life)
- QALY = number of years lived x utility weight
  - One healthy year of life = 1 QALY (1 year of life × 1 utility weight)
  - One year of life lived in situation with illness/disability (e.g. chronic pain) = 1 year × 0.5 (utility weight) = 0.5 QALYs
  - *Death:* assigned value of 0 QALYs

NOTES
NON-PARAMETRIC TESTS

NON-PARAMETRIC TESTS

- For data that is assumed to not be distributed normally
- For nominal/ordinal level variables

CHI-SQUARED TEST

osms.it/chi-squared_test

- Chi-square (X^2) goodness-of-fit test
- Test compares categorical variables
  - Assesses for significant association
- Examines whether collected data is significantly different than theoretical model
  - How "good is the fit" between data, what is expected
- *Null hypothesis*: no significant difference between theorized/expected, observed ratios
  - X^2 = sum of [(observed – expected)^2/ expected]
- Use X^2 table to find critical X^2
  - Adjusted for degrees of freedom [n – 1], at selected p-value
- Accept null hypothesis if X^2 < critical X^2

CHI-SQUARE TEST OF INDEPENDENCE

- For analysis of contingency tables (or crosstabs tables)
- Investigates whether two/more categorical variables are statistically significant
- Used for multiple variables
- Degrees of freedom = (# of rows – 1) x (# of columns – 1)
- Requires > 5 data points in all cells of table; whole numbers
- Higher X^2 results in lower p-value

FISHER'S EXACT TEST

osms.it/Fisher_exact_test

- Variant of chi-square test
  - Used with small sample size
- Used to determine exact probability of association between two categorical variables (i.e. significance of association [contingency] between classifications)
  - Use for 2 x 2 contingency tables (< 5 in a cell)
  - p-values calculated exactly
  - $p < 0.05$, unlikely to be random association

KAPLAN-MEIER SURVIVAL ANALYSIS

osms.it/Kaplan-Meier_survival_analysis

- Estimates survival from lifetime data; measures fraction of survivors over treatment time; simplest method of computing survival over time
  - Plot of percent survival versus time; generated from status at last observation, time to event
  - Large sample size → approaches population effect
  - Accounts for censored data; withdrawn from study, lost to follow-up; alive at last follow-up (i.e. right-censoring—data above a certain value, but otherwise unknown)
  - Limited capacity to estimate survival adjusted for covariates

KAPPA COEFFICIENT

osms.it/kappa-coefficient

- AKA Cohen's kappa coefficient
- Measure of inter-rater agreement
- Compares ability of different raters to classify categorical variables
- *Interobserver agreement*: accounts for agreement that occurs by chance, when raters measure same thing, using same observation method
- Calculated from observed, expected frequencies from diagonal of contingency table
- If kappa = 1
  - Agreement is perfect
- If kappa = 0
  - Agreement is no better than if agreement happened by chance
- Example for interpreting agreement based on kappa coefficient
  - *None:* < 0
  - *Fair:* 0.20–0.40
  - *Moderate:* 0.40–0.60
  - *Good:* 0.60–0.80
  - *Very good:* 0.80–1.00

MANN-WHITNEY U TEST

osms.it/Mann-Whitney_u_test

- Nonparametric test equivalent to unpaired t-test
- Compares differences between two unpaired groups that are not normally distributed
  - Uses number ranks rather than raw data
  - Provides p-value indicating whether or not groups are significantly different from each other ($p < 0.05$; unlikely to happen by chance)

SPEARMAN'S RANK CORRELATION COEFFICIENT

osms.it/Spearmans-rank-correlation-coefficient

- Spearman's rho (ρ)
- Non-parametric equivalent of Pearson's correlation coefficient
- Measure strength, direction of monotonic association between two ranked variables
  - Monotonic association means variables increase together (i.e. as value of one variable increases value of other variable increases also/as value of one variable increases, other variable value will decrease)
  - Does not have to be linear, but must be entirely increasing/entirely decreasing (may include plateaus)
  - Use for continuous/discrete ordinal, interval, ratio variables
- Sign indicates direction of association
  - x and y increasing → +ve ρ; x and y decreasing → –ve ρ
- ρ increases as correlation approaches perfect monotone relationship between variables
- Two formulas
  - One for when there are no tied ranks
  - One for tied ranks
- Use critical values (r_s) from Spearman's rank coefficient tables to determine significance of r (Spearman's coefficient of sample)

NOTES
PARAMETRIC TESTS

PARAMETRIC TESTS

- ANOVA, t-tests
- Use for following data
  - Randomly selected samples
  - Independent observations
  - Population standard deviations (SDs) are same
  - Data distributed normally/approximately normally

ANOVA

osms.it/one-way_ANOVA
osms.it/two-way_ANOVA
osms.it/repeated-measures_ANOVA

- AKA analysis of variance
- Determines differences between > two samples
  - Measures differences among means
- F-ratio (F statistic)
  - F = (variance between groups) / (variance within each group)
- Computer program calculates p-value from F; use F to accept/reject null hypothesis
  - F approx. = 1; p large; accept null hypothesis
  - F large → p small (alpha set at 0.05 significant → reject null hypothesis)
- Assumptions
  - Samples drawn randomly; sample groups have homogeneity of variance (i.e. from same population; interval, ratio data)

1-way ANOVA

- Between groups design
- One independent variable
  - May have multiple levels (e.g. drug A effect vs. drug B vs. placebo on specified outcome)

Factorial ANOVAs

- Factorial designs
- Two-way, three-way, four-way ANOVA, more (two, three, four, etc. independent variables)

Single-factor repeated measures ANOVA

- ANOVAs involving repeated measures/ within groups/subjects
- One independent variable with multiple levels tested within one subject group (e.g. drug A vs. drug B vs. placebo tested within same individuals at different times)
- ↓ variation effect between sample groups

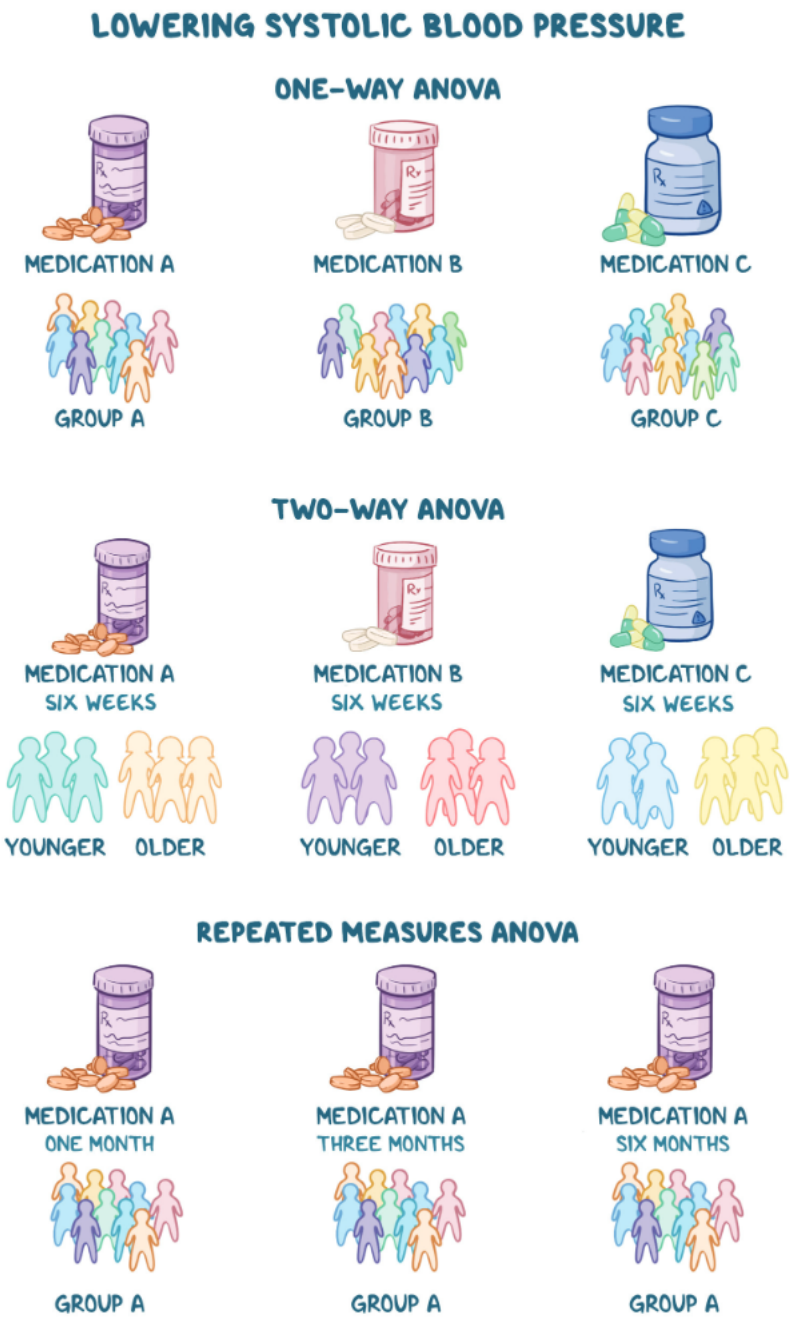


Figure 10.1 Examples demonstrating a one-way, two-way, and repeated measures ANOVA. The one-way ANOVA has one independent variable (medication type) with multiple levels (medications A, B, and C). The two-way ANOVA looks at two independent variables (medication type and age category) that each have multiple groups (medications A, B, and C; younger and older). The repeated measures ANOVA follows the same group of people over a period of time to measure the effects of the same medication over time. In this case, the independent variable is time, divided into three groups (one month, three months, and six months), and the dependent variable is systolic blood pressure.

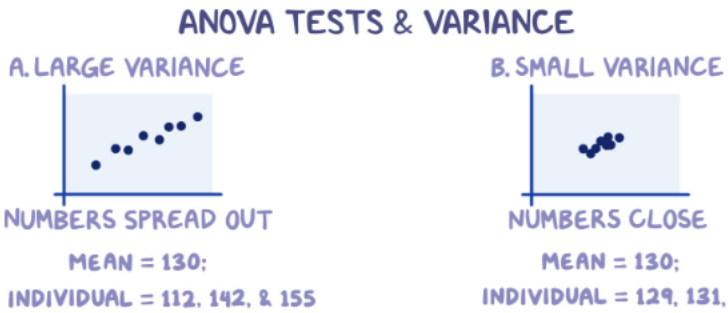


Figure 10.2 All ANOVA tests assume that the groups have equal variance. A large variance means that the numbers are very spread out from the mean; a small variance means that the numbers are very close to the mean. Variances between groups are considered unequal when the variance of one group is greater than twice the variance of the other group.

CORRELATION

osms.it/correlation

- Investigates relationships between variables; determines strength, type (positive/negative) relationship
- *Correlation coefficient*: r ($-1 > r < +1$)
  - *Perfect positive correlation:* $r = +1$
  - *Perfect negative correlation:* $r = -1$
  - *No correlation:* $r = 0$
  - *Strong correlation:* $r > 0.5 < -0.5$
  - *Weak correlation:* $0 < r < 0.5$, or $0 > r > -0.5$
- *Pearson product-moment coefficient*: interval/ratio data; calculates linear relationship degree between two variables
- *Confidence interval (CI)*: population based on correlation coefficient
  - Indicates range within population correlation coefficient lies
- P-value for correlation coefficient based on null hypothesis
  - I.e. if true ($p > 0.05$), no correlation between variables
- Coefficient of determination: r^2 or R^2 ($0 < R^2 < 1$)
  - Fraction of variation of variable of interest (x axis) due to another variable of interest (y axis)
  - Remaining proportion due to natural variability
  - Low R^2 may indicate poor linear relationship, may be strong nonlinear relationship
- *Eta-squared* (η^2): analogous to R^2 for ANOVA
- Correlation ≠ causation, consider
  - How strong is association?
  - Does effect always follow cause?
  - Is there a dose response?
  - Relationship biologically plausible, coherent?
  - Consistent finding?
  - Other factors involved?
  - Good experimental evidence?
  - Analogous examples?

Figure 10.3 Scatterplots are used to plot measurements, with one measured variable on each axis. Each data point represents one individual. A trend line is drawn to best represent the collection of data points on the plot, with roughly half the points above the line and the other half below the line. A perfect positive or negative correlation means that the trend line passes through every single data point.

HYPOTHESIS TESTING

osms.it/hypothesis-testing

- Calculating sample size required to test hypothesis
- Equations used for calculating power can also be used to calculate sample size for a predefined alpha (0.05)
- Requires knowledge of
  - Clinically important effect size (larger sample size needed to detect smaller effects)
  - Surrogate endpoint use rather than direct outcome
  - Desired power; alpha (if not 0.05); confidence interval
  - Statistical tests to be used
  - Data lost to follow-up
  - Test group SD; population of interest expected frequency within test group
- Statistician's advice
  - Optimize sample size, avoid underpowered studies, enable valid data interpretation

LINEAR REGRESSION

osms.it/linear-regression

- *Simple linear regression*: assumes linear relationship; slope ≠ 0; data points close to line
- Examine weight of two variables' (x, y) effects; predict effects of x on y
- Fit best straight line to x, y plot of data
  - Equation: $y = bx + a$ (x and y are independent variables; b = slope of line (regression coefficient); a = intercept)
- 95% CI for slope range; larger sample → narrower CI; if range does not include zero → real correlation suggested
- p-value for null hypothesis
  - No linear correlation (i.e. slope = 0; $p < 0.05$ → real correlation suggested)

OTHER REGRESSION ANALYSES

- Multiple linear regression
  - Examines effects of more than one variable on y
- Multiple nonlinear regression
  - Examines correlations among nonlinear data, more than one independent variable

- Logistic regression
  - Predicts likelihood of categorical event in presence of multiple independent variables

LOGISTIC REGRESSION

osms.it/logistic-regression

- *Predictive analysis:* describes relationship between binary dependent variable (i.e. takes one of two values), multiple independent variables
- Assumptions
  - *Dichotomous outcome* (e.g. yes/no; present/absent; dead/alive)
  - *No outliers:* assess using z scores
  - *No intercorrelations*: assess using correlation matrix
- May use logit (assumes log distribution of event's probability)/probit (model assumes normal distribution)
- *Rule of 10*: stable values if based on minimum of 10 observations per independent variable
- *Regression coefficients*: indicate contribution of individual independent variables; odds ratios
- Tests to assess significance of independent variable
  - Likelihood ratio test; Wald test
- *Bayesian inference*: prior (known) distributions for regression coefficients; conjugate prior; automatic software (e.g. OpenBUGS, JAGS to simulate priors)

TYPE I & TYPE II ERRORS

osms.it/type-I-and-type-II-errors

POWER

- Refers to test probability correctly rejecting false null hypothesis
- *Power*: (1 – beta)
  - Likelihood that statistically non-significant result is correct (i.e. not false negative—type II error)
- Medical research
  - Power typically set at 0.80
- Increasing power
  - ↓ type II error chance; ↑ type I error chance
- Power increases when ↑ sample size, ↓ SD, ↑ effect size

EFFECT SIZE

- Relationship strength between variables
- Statistical significance does not necessarily indicate clinical significance
- Random variation (SD) may ↓ differences between outcomes of interest between hypothesis' test groups
- $$ES = \frac{\overline{X}_1 - \overline{X}_2}{SD}$$
  - ES is effect size
  - X_1 is the mean for Group 1
  - X_2 is the mean for Group 2
  - SD is the standard deviation from either group

- Adjust for variation in test groups with Cohen's d (assumes each group's SD is same)
  - Cohen's d = (mean 1 – mean 2)/SD
  - 0.2 = small effect size
  - 0.5 = medium effect size
  - ≥ 0.8 = large effect size

SAMPLE SIZE

- Smaller sample size
  - ↑ sampling error chance
  - Lower power
  - ↑ type II error chance (false negative)

BAYESIAN THINKING

- Relates p-value to context
  - Can involve complex mathematics
- Measures event probability given incomplete information
- Joint distribution between given information (usually probability density), experimental results

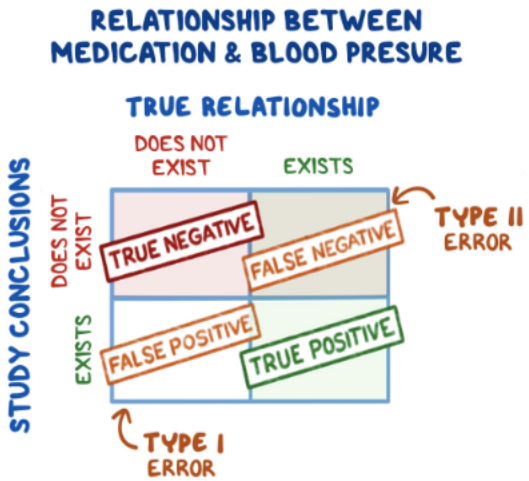


Figure 10.4 A Type I error occurs when no true relationship exists between two variables, but the study concludes there is one; a type II error occurs when there is a true relationship between two variables, but the study concludes there is no relationship.

NOTES
STATISTICAL PROBABILITY DISTRIBUTIONS

NORMAL DISTRIBUTION & Z-SCORES

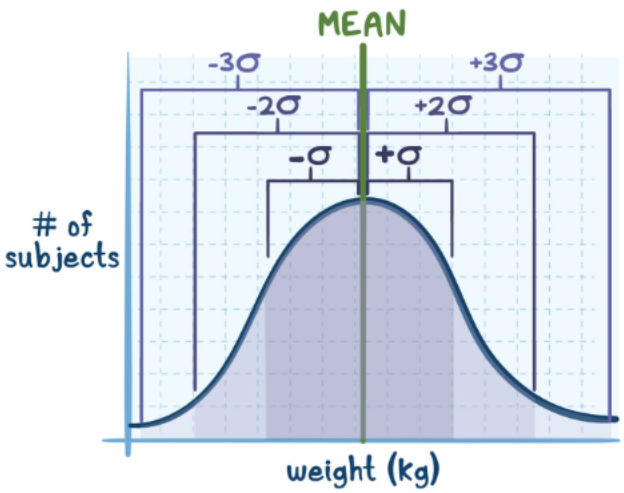
osms.it/normal-distributions-z_scores

NORMAL DISTRIBUTION

- Data grouped around central value, no left/ right bias, in "bell curve" shape
- Probability distribution for normal random variable x
- $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(\frac{1}{2})[(x-\mu)/\sigma)^2}$
- μ = mean of normal random variable x
- σ = standard deviation
- π = 3.1416 ...
- e = 2.71828 ...
- *Normal distribution:* $\mu = 0$, $\sigma = 1$

Z-SCORES

- Standardized score
- Uses data set mean, standard deviation to determine measurement location
  - Represents deviation from mean
- Expressed in standard deviations
- Sample z-score for measurement x
- $z = \frac{x - \overline{u}}{\sigma}$
- μ = population mean
- σ = standard deviation



BELL CURVE

1 standard deviation:
68% = +/- σ

2 standard deviations:
95% = +/- 2σ

3 standard deviations:
99% = +/- 3σ

STANDARD ERROR OF THE MEAN

osms.it/standard-error-of-mean

- AKA SEM, standard deviation
- $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
- σ = standard deviation
- n = sample size

PAIRED T-TESTS

osms.it/paired-t-test

- Statistical hypothesis test (parametric)
- Determines if two groups are statistically different (compares two groups' means)
- Groups can occur naturally (e.g. smokers compared to non-smokers)/groups can be created experimentally (e.g. control group compared to treatment group)
- t = difference between means / variance/sample size
- = sample mean - population mean / sample standard error of the mean
- $\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
- x_1 = mean of sample 1
- x_2 = mean of sample 2
- n_1 = sample size of sample 1
- n_2 = sample size of sample 2
- s_1^2 = variance of sample 1 = $\frac{\sum(x_1 - \bar{x}_1)^2}{n_1 - 1}$
- s_2^2 = variance of sample 2 = $\frac{\sum(x_2 - \bar{x}_2)^2}{n_1 - 1}$

ONE-TAILED & TWO-TAILED TESTS

osms.it/one-tailed-two-tailed-tests

- *Tails*: ends of probability curve
- Alternative (research) hypothesis proposes groups under investigation are different in some way/relationship between them exists

ONE-TAILED TESTS

- Alternative hypothesis is directional (i.e. specifies direction of difference/relationship)
  - Extreme values of one of distribution tails are of interest given difference type/ expected relationship (solid theoretical

basis required for one-tailed test)

- Alternative hypothesis predicts relationship between groups either positive/negative (e.g. Group A will score higher on particular test than Group B)

TWO-TAILED TESTS

- Alternative hypothesis is non-directional (i.e. non-specified direction of difference/ relationship)
  - Extreme values on either tail of sampling distribution support null hypothesis rejection (e.g. Group A scores will be different than Group B)

NOTES
STUDY DESIGN

SAMPLING

osms.it/sampling

- Selection of individuals for study from specific population
- Aims to represent, estimate characteristics of that population

PLACEBO EFFECT & MASKING

osms.it/placebo-effect-and-masking

WHAT IS THE PLACEBO EFFECT?

- Refers to situation where study participant's belief in treatment brings about positive effect
  - E.g. individuals given placebo drug tend to report improvements even when treatment has no real effect
- Placebos can be affected by study participant's psychological responses to context in which treatment is taking place
- Placebo can be drug/pharmacologically inactive substance indistinguishable from an active treatment/can be based on any expectation the person may have about intervention under study
- Useful in studying rate of side effects, reactions to drug

WHAT IS MASKING?

- Subjects and/or investigators are unaware of treatment assignment
  - *Single blind:* subjects are unaware of treatment assignment
  - *Double blind:* subjects, investigators are unaware of treatment assignment
  - *Triple blind:* treatment administrator unaware of treatment assignment

CASE-CONTROL STUDY

osms.it/case-control_study

- Study that determines potential risk factors in individuals with condition
- May rely on individual recall, past medical history, autopsy
- *Example:* Percentage of people who gave birth to child with condition A who had previously taken drug B during pregnancy
  - All children either do or do not have condition A
  - We assess whether they did/did not

take drug B during pregnancy

Pros

- More easily examines rare diseases than prospective studies; less expensive and time-consuming
- Individuals not exposed to possible risk factors
- Past medical history used to determine potential multiple risk factors

Cons

- Potential problems matching cases and controls
  - E.g. study may be influenced by characteristics not being studied (confounding variables)
- Potentially biased (relies on individual recall)
  - E.g. study candidates may emphasize potential risk factors rather than controls

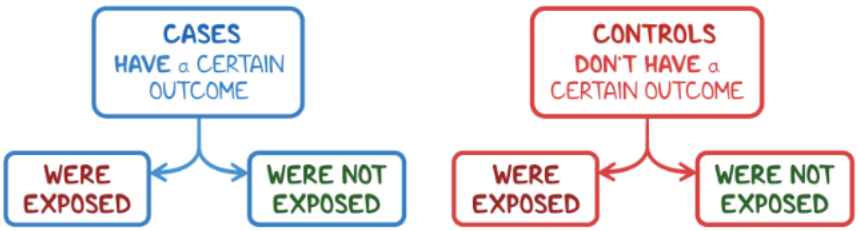


Figure 12.1 Case-control study design.

COHORT STUDY

osms.it/cohort-study

- Measures disease within group of individuals (cohort) over period of time
- Focuses on disease development
- *Two types:* prospective cohort, retrospective cohort

PROSPECTIVE COHORT STUDY

- AKA longitudinal, concurrent cohort study
- Results not known until after intervention
- Used to follow up on people who received treatment/were exposed to risk factors
- Laboratory tests often used as surrogate markers – for example, increase in hemoglobin immediately after blood transfusion assumed to mean that transfusion was effective
- *Example:* RSV rates of premature birth cohorts

Pros

- Easier to conduct than randomized controlled studies
- Useful information on risk
- Matching decreases influence of confounding variables

Cons

- Expensive, time-consuming
- Follow-up with people over time can be difficult; subjects may be lost

RETROSPECTIVE COHORT (HISTORICAL COHORT, NONCONCURRENT PROSPECTIVE) STUDY

- Same prospective cohort study design but uses past data to determine future time frame; study and obtention of results faster
- Use pre-existing population to decrease study duration
- Can be conducted relatively quickly, inexpensively
  - E.g.mortality rates according to duration of smoking

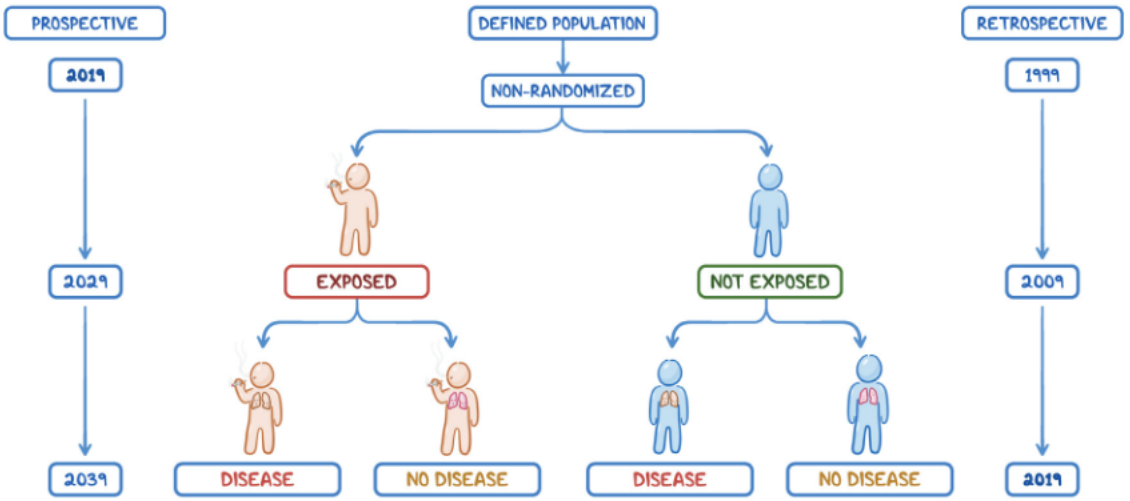


Figure 12.2 Design of prospective and retrospective cohort studies with hypothetical time frames. Exposed = smokers, not exposed = non-smokers, disease = lung cancer.

CROSS-SECTIONAL STUDY

osms.it/cross-sectional_study

- Study that observes a group of people at one point in time
- Examines relationship between an exposure (variable), disease being investigated
- *Example:* the relationship between endometrial cancer, hormone replacement therapy (HRT)

Pros

- Less time-consuming, expensive than longitudinal studies, as individual follow-up not necessary
- Good for establishing overall association between exposure and disease
- Can establish disease prevalence (number of individuals with particular disease in their lifetime)

Cons

- Establishes disease prevalence but not incidence (percentage of individuals who may develop a particular disease within a year)
- Does not establish temporal relationship between exposure and disease
- Potentially biased if surveys used
- *Retrospective studies:* data quality may be compromised due to poor recall/"recall bias," where people are more likely to recall certain events

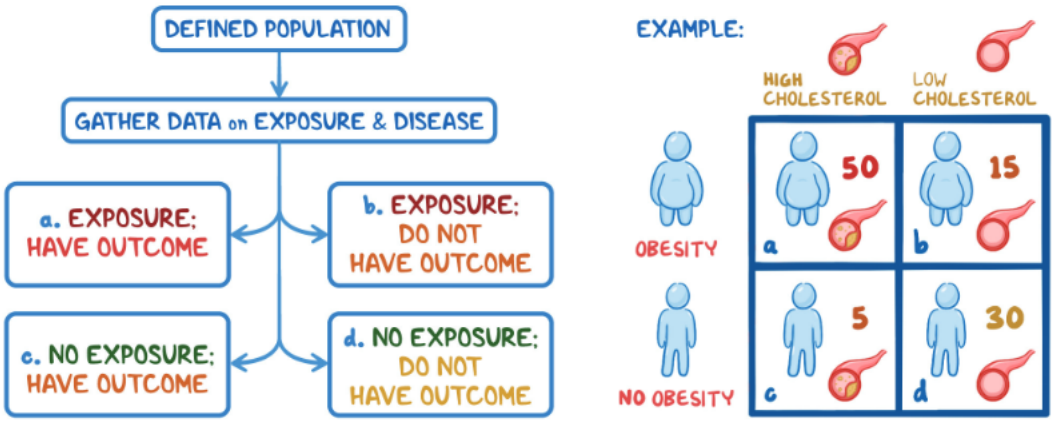


Figure 12.3 Design of a cross-sectional (prevalence) study. *Example:* obesity is the exposure, and high cholesterol is the outcome.

ECOLOGIC STUDY

osms.it/ecologic-study

- Observes at least one variable
  - Exposure/outcome
- Measured at group level
- At least one comparison group, disease occurrence compared between groups
- Often used to make large-scale comparisons

- Examples
  - Rate of cancer occurrence in one population
  - Average sunlight exposure at different geographical locations
  - Comparing per capita dietary fat consumption, cardiovascular disease mortality
  - Disease occurrence compared between groups

RANDOMIZED CONTROL TRIAL (RCT)

osms.it/randomized-control-trial

- Examines effectiveness of intervention (e.g. medications, treatment protocols)
- *Three features:* randomization, control, manipulation
- Considered gold standard of experimental research, identifying cause-and-effect relationships
- Study participants randomly assigned either experimental group or control group
- *Example:* Effects of drug A versus drug B on hypercholesterolemia in individuals with type 2 diabetes mellitus

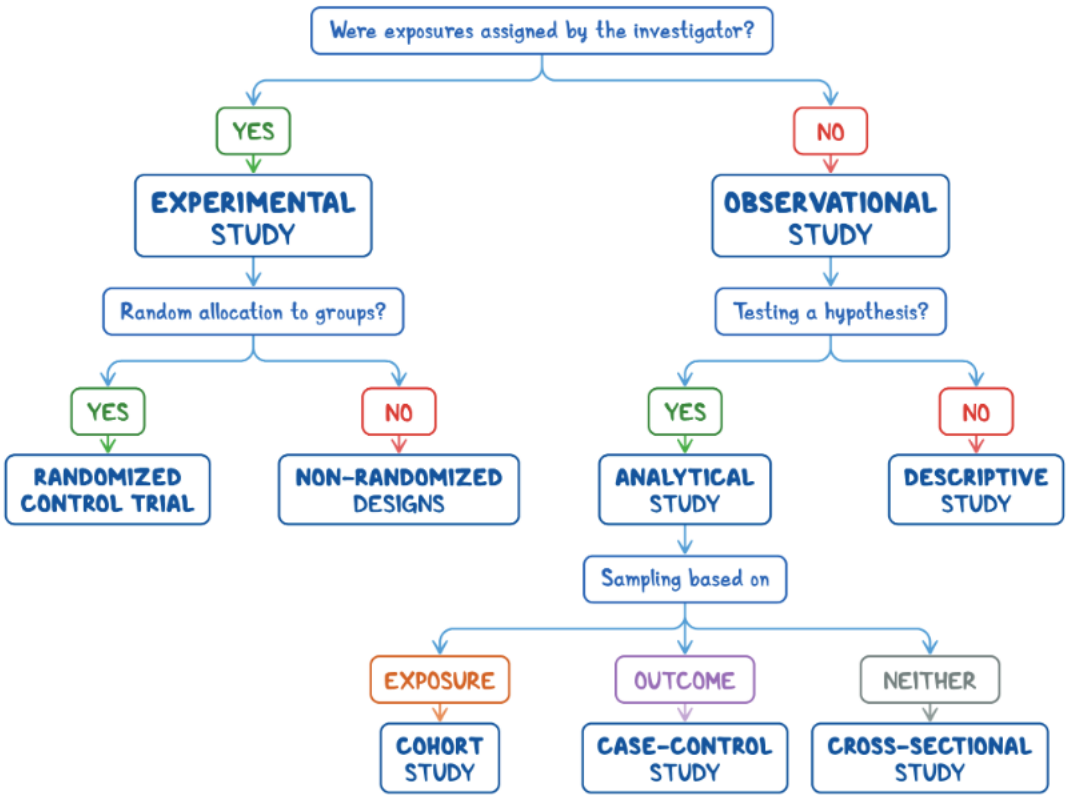


Figure 12.4 A summary flowchart of the different types of study designs.

NOTES
TESTING

SENSITIVITY (SN) & SPECIFICITY (SF)

osms.it/sensitivity-specificity

- Validity measure; concerned with how close test's result is to truth (i.e. did test/ instrument measure what it is intended to measure?)
  - No perfect test → some miscalculation degree inevitable (i.e. healthy individual tests positive for disease → false positive; sick individual tests negative → false negative)
  - *Sn*, *Sp*: complementary test characteristic measures must be used together

SENSITIVITY

- Population proportion who test positive for disease, have disease
- AKA true positive rate
- Highly sensitive test with positive result identifies people who are truly diseased (true positives), some healthy people (false positives)
- *Sensitivity:* proportion containing all truly positive, false positives
- Can assume two things
  - Test with high sensitivity is negative, individual must be healthy → rule out disease
  - Test with high sensitivity is positive, individual may/may not have disease (ensure lack of false positive; further testing required)
  - High sensitivity negative test → useful for ruling-out disease

SPECIFICITY

- Population proportion tests negative for disease, free of disease
  - AKA true negative rate
- Highly specific test with negative result
- Identifies all people who are truly free of disease (true negatives), some sick people (false negatives)
- *Specificity:* proportion containing all truly negative, false negatives; two things assumed
  - Test with high specificity positive → confirm disease
  - Test with high specificity negative → individual may/may not have disease (ensure not false negative; further testing required)
- Positive test with high specificity → useful disease confirmation

CUTOFF POINT

- *For continuous variables:* sensitivity, specificity may overlap → midpoint usually sought (avoids misclassification)
- Cutoff point needed to distinguish between normal/healthy, abnormal/unhealthy results

High cutoff point

- *Highly specific:* low false positives
  - Everyone categorized as abnormal has disease
- *Poorly sensible:* high false negatives
  - Not everyone categorized as normal is free of disease

- I.e. previous hypertension definition stated 140/90mmHg as cutoff point
  - *Highly specific:* everyone categorized as abnormal has disease
  - *Poorly sensitive:* not everyone categorized as normal is free of disease

Low cutoff point

- *Poorly specific:* high false positives
  - Not everyone categorized as abnormal has disease
- *Highly sensitive:* low false negatives
  - Everyone categorized as normal is free of disease
- I.e. new hypertension definition states 120/80mmHg as cutoff point
  - *Poorly specific:* not everyone categorized as abnormal has diseases
  - *Highly sensitive:* everyone categorized as normal is free of disease

Cutoff point determined by test's purpose

- Screening test
  - Needs to detect all possible diseased → low cutoff point → highly sensitive → low false negatives
- Confirmatory test
  - Need to be sure of disease presence → high cutoff point → highly specific → low false positives

SEQUENTIAL & SIMULTANEOUS TESTING

Sequential testing

- AKA two-stage testing
- Consecutive tests performed with different characteristics → obtain more specific results
  - Perform first test → positive → perform second test → positive → disease likely present
  - Perform first test → negative → disease not likely present
- Similar to "double checking" results
- First test often easier/cheaper/less invasive than second test
- Sensitivity, specificity calculations must include both tests' characteristics
- *Net sensitivity:* proportion of true cases that test positive on both first, second test
  - First test sensitivity x second test sensitivity
- *Net specificity:* proportion of healthy people that test negative on either first, second test
  - (First test specificity + second test specificity) - (first test specificity * second test specificity)

Simultaneous testing

- Two tests with different characteristics performed at same time → more sensitive results
  - Simultaneous testing: three groups of people
  - People detected only by Test A
  - People detected only by Test B
  - People detected by both Test A and Test B
  - Pools all possibly relevant information → more sensitive results
- Sensitivity, specificity calculations must include both tests' characteristics
- *Net sensitivity:* proportion of true cases that test positive on either test A or B
  - (Test A sensitivity + Test B sensitivity) - (Test A sensitivity x Test B sensitivity)
- *Net specificity:* proportion of healthy people that test negative on both tests A and B
  - Test A specificity x Test B specificity

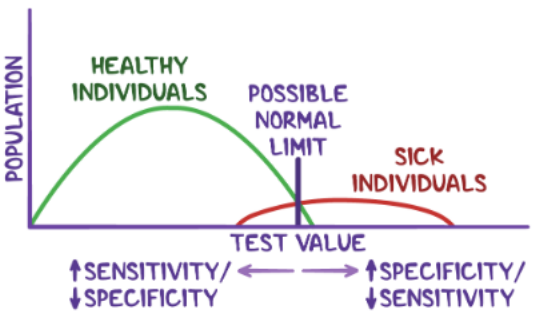


Figure 13.1 Illustration showing how sensitivity and specificity are affected by moving the cut-off point.

SENSITIVITY & SPECIFICITY

		DISEASE		TESTING TOTAL	MEASURES
		PRESENT	ABSENT		
TEST	POSITIVE	True positives (TP)	False positives (FP)	Total positive test (TP + FP)	Positive Predictive Value (PPV) = true positives / total positives
	NEGATIVE	False negatives (FN)	True negatives (TN)	Total negative test (FN + TN)	Negative Predictive Value (NPV) = true negatives / total negatives
TOTAL DISEASE		Total disease (TP + FN)	Total no disease (FP + TN)	Total population (TP + FP + TN + FN)	Muscles, liver
MEASURES		Sensitivity = true positives / total disease	Specificity = true negatives / total no disease	Prevalence = Total disease / total population	

POSITIVE & NEGATIVE PREDICTIVE VALUE

osms.it/positive-negative-predictive-value

- *PPV:* probability that if test is positive, person has disease
  - Divide true positives, total positive test number
- *NPV:* probability that if test is negative, person is free of disease
  - Divide true negatives, total negative test number

- Both measures directly influenced by prevalence, test specificity
  - *High prevalence:* more likely that person has disease → ↑ PPV
  - *Low prevalence:* less likely that person has disease → ↑ NPV
  - *Low prevalence:* need a good test in confirming disease (high specificity) → ↑ PPV

TEST PRECISION & ACCURACY

osms.it/test-precision-accuracy

- Both concerned with how likely test to be reproduced → return results close to truth
  - Neither measuring devices nor people perfect → affects test precision, accuracy
- *Test precision:* how repeatable test results are over time, regardless of result accuracy
  - *High precision test:* consistently deliver similar results, regardless of whether true/not
- *Test accuracy:* how true test results are, regardless of test repeatability
  - *High accuracy test:* gives correct results; cannot always be reproduced

- Comparing test precision, accuracy
  - Oximeter consistently (precisely) reports true pO_2 (accurately)
  - Oximeter consistently (precisely) reports pO_2 20% lower than truth (not accurate)
  - Oximeter inconsistently (not precise) reports true pO_2 (accurate)
  - Oximeter inconsistently (not precise) reports pO_2 20% lower than truth (not accurate)