

We are Right 95% of the Time: Fundamentals of Biostatistics for Healthcare Professionals

Daniel Majerczyk, PharmD, BCPS, BC-ADM, CACP

Associate Professor of Clinical Sciences
Roosevelt University College of Science, Health
and Pharmacy
dmajerczyk@roosevelt.edu

Ashley Stefanski, PharmD

Clinical Instructor, Academic Fellow
Roosevelt University College of Science, Health
and Pharmacy
astefans@roosevelt.edu



Objectives:

1. Review basic statistical principles, such as the different types of variables and statistical tests to critically evaluate medical literature
2. Interpret basic biostatistical principles, such as confidence intervals, ratio data (risk ratio, hazard ratio), the number needed to treat, and number needed to harm
3. Apply the information learned to a piece of medical literature



Conflict of Interest

- The speakers have no potential conflicts of interest or financial disclosures



Poll Question

On a scale of 1-5, how would you rate your skills in interpreting medical literature?

- A. 1 - Poor
- B. 2 - Fair
- C. 3 - Average
- D. 4 - Good
- E. 5 - Excellent



Basic Biostatistic Principles

Independent and Dependent Variables

- **Independent variables:**

- Set (manipulated/changed) by researcher
- Controlled by researchers to determine whether it has an effect on the dependent variable

- **Dependent variables (outcomes):**

- Variables that will DEPEND upon the independent variables

Independent and Dependent Variables

Independent variables – things that the researchers can change/set

Medication	Medication dose	Placebo	Patients
------------	-----------------	---------	----------



Dependent variables (outcomes)

Heart failure progression	A1c	BP (blood pressure)	LDL level
---------------------------	-----	---------------------	-----------

Poll Question – Word Cloud

In the following example, identify which variable is the dependent variable:

Researchers are designing a study that is analyzing weight loss after starting a new exercise program.

Types of Statistical Variables - Categorical Variables

CATEGORICAL ↓ Fits into a limited number of categories	<u>NOMINAL DATA</u>	CATEGORIES ARE IN AN ARBITRARY ORDER	
		ORDER OF CATEGORIES DOES NOT MATTER	Gender, Ethnicity Marital status, Mortality
	<u>ORDINAL DATA</u>	CATEGORIES ARE ARRANGED IN A LOGICAL ORDER, BUT THE DIFFERENCE BETWEEN CATEGORIES IS NOT EQUAL	
		ORDER OF CATEGORIES DOES MATTER	NYHA Functional Class I – IV Pain scale 0 - 10

Quantitative Variables – Continuous Variables

CONTINUOUS ↓ Provided by some type of measurement which has unlimited options of continuous values	<u>RATIO SCALE</u>	EQUAL DIFFERENCE BETWEEN VALUES WITH A MEANINGFUL ZERO	0 = NONE • Age • Height • Weight • Time • BP (Blood Pressure)
	<u>INTERVAL SCALE</u>	EQUAL DIFFERENCE BETWEEN VALUES, NO MEANINGFUL ZERO	0 ≠ NONE • Temp. in Celsius • Temp. in Fahrenheit

Poll Question

The primary outcome in a study of pediatric patients with mild persistent asthma comparing the efficacy of fluticasone versus budesonide is change from baseline in pulmonary function tests. Which describes the type of data that will be measured?

- A. Continuous - Ratio
- B. Continuous - Interval
- C. Nominal
- D. Ordinal

Poll Question

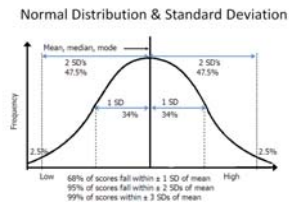
Baseline characteristics in a trial included age, sex, and smoking status. What type of data do these values represent, respectively?

- A. Interval, nominal, ordinal
- B. Interval, nominal, nominal
- C. Ratio, nominal, nominal
- D. Ratio, ordinal, nominal

Descriptive Statistics – Measures of Central Tendency

$$\text{Mean} = \frac{\text{Sum of all values}}{\text{Number of values}}$$

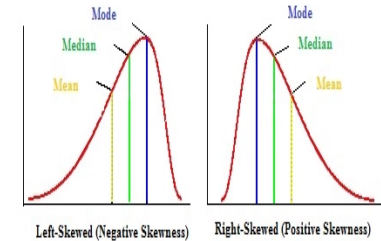
- **Mean**
 - Average value
 - Use when the data are not skewed
 - Used for continuous data
- **Median**
 - Value in the middle of the ranked list
 - 50th percentile
 - Use when continuous data are skewed
 - Preferred for ordinal data
- **Mode**
 - Value that occurs most frequently in a set of data
 - Preferred for nominal data



http://www.bradfordresearch.nhs.uk/uploads/downloads/step-by-step-guide/Data_Analysis_for_Researchers.pdf
Gaddis ML et al. *Ann Emerg Med*. 1990; 19:1462-8.

Skewness

- Data that is not normally distributed are skewed
- Data skewed to the right
 - Positive skew
- Data skewed to the left
 - Negative skew
- Direction of the skew refers to the direction of the longer tail



<https://www.statisticshowto.datasciencecentral.com/probability-and-statistics/skewed-distribution/>
Gaddis ML et al. *Ann Emerg Med*. 1990; 19:1462-8.

Poll Question

True or false: Median should only be used for continuous, normally distributed data

- A. True
- B. False

Choosing The Appropriate Statistical Test

- **Step 1**
 - Identify scale of measurement
 - Nominal, ordinal, continuous
- **Step 2**
 - Establish distribution
 - Parametric vs. nonparametric
 - Normal vs. non-normal distribution
- **Step 3:**
 - Distinguish groups
 - Independent vs. dependent
 - Quantity of groups

Statistical Tests

	Independent		Dependent	
Number of groups	2 groups	≥2 groups	2 groups (either crossover or pre/post treatment)	Multiple measures in ≥2 groups
Continuous Data	t-test	ANOVA	Paired t-test	Repeated Measures ANOVA
Ordinal Data	Mann-Whitney U Test (Wilcoxon Rank Sum)	Kruskal-Wallis	Wilcoxon Signed Rank Test	Friedman
Nominal Data	Chi-square or Fisher's Exact Test	Chi-square or Fisher's Exact Test	McNemar	Cochrane Q

Poll Question

Which data does the Mann-Whitney U test compare?

- A. Parametric and ordinal
- B. Nonparametric and nominal
- C. Parametric and nominal
- D. Nonparametric and ordinal

Poll Question

A hospital pharmacy director wants to assess pharmacists' confidence level before and after the introduction of a new rapid response simulation activity. Confidence was assessed on a 5-point Likert scale. What is the best statistical test to compare this data?

- A. Chi-square test
- B. McNemar's test
- C. Student's t-test
- D. Wilcoxon signed rank test

Poll Question

Researchers are assessing the effects of the treatment of heart failure on functional class in three groups of adults after 12 months of treatment. Investigators wanted to assess functional class with the New York Heart Association (NYHA) functional classification, and to compare the patient classification after 12 months of treatment.

See next slide for question

Poll Question

Which statistical test is most appropriate to assess differences in functional classification between the groups?

- A. Analysis of variance (ANOVA)
- B. Wilcoxon signed-rank test
- C. Kruskal-Wallis test
- D. Mann-Whitney U test

Determining Significance

Hypothesis Testing for Significance

• H_0 & H_a

- H_0 - null hypothesis
 - Accepted when no statistical difference is detected between intervention and control group(s)
 - In biostatistics;
 - Generally, H_0 is accepted when p-value is ≥ 0.05
 - This is usually what the researcher is trying to disprove or reject
- H_a - alternative hypothesis
 - Opposite of the null hypothesis; when H_0 is rejected, H_a is accepted
 - There is a significant difference between the intervention and control group(s)
 - In biostatistics;
 - This is usually what the researcher is trying to prove or accept

The p-value (α) and What It Represents

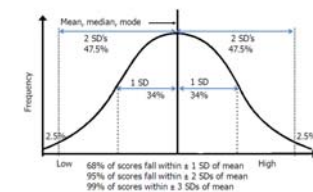
• p-value

- Likelihood of a Type 1 Error
 - A chance that what you found in your study is actually different from reality

- The p - value is also referred to as alpha (α)

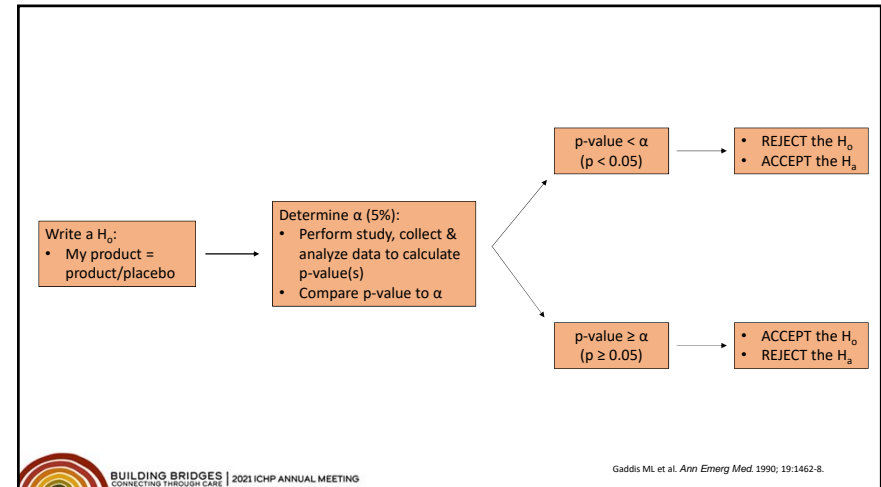
- α
 - Correlates with the values in the tails when data has a normal distribution

Normal Distribution & Standard Deviation



The p-value (α) and What It Represents

- If α is set at 0.05 and the p-value $< \alpha$ ($p < 0.05$):
 - The null hypothesis is **REJECTED** and thus the results **are** statistically significant
 - Alternative hypothesis is **ACCEPTED**
- If the p-value $\geq \alpha$:
 - The study has **failed to reject** the null hypothesis and thus the results **are NOT** statistically significant
 - Alternative hypothesis is **REJECTED**



The p-value (α) and What It Represents

- **Type I Error**
 - Detecting a difference when in reality one does not exist
 - A false positive
- The probability of making a Type I Error is determined by α and it relates to the Confidence Interval (CI)
 - **Confidence Interval = $1 - \alpha$ (Type I Error)**
 - If $\alpha = 0.05$
 - $1 - 0.05 = 0.95$
 - If your results yielded a $p < 0.05$ = Statistically Significant
 - The probability of making a Type I Error is $< 5\%$
 - You are 95% confident that your result is correct and NOT due to chance

The Beta (β) Value and What It Represents

- β = probability of Type II Error
- **Type II Error**
 - % or chance of NOT detecting a difference when one actually does exist
 - False negative
- β is typically set at 0.1 or 0.2
 - The risk of a Type II Error is 10 or 20 times in 100
 - This relates to the study POWER
- **Power = $1 - \beta$**
 - The probability that a test will REJECT the null hypothesis correctly or the ability to detect a significant difference if one exists
 - As power \uparrow , the chance of a Type II Error \downarrow
 - The higher β is (or the higher the probability of Type 2 Error), the less power your study will have

		Researcher Interpretation	
Reality		H_0 Accepted/ H_a Rejected	H_0 Rejected/ H_a Accepted
	H_0 is true (No difference)	Correct conclusion ($p \geq 0.05$)	Type I Error
	H_0 is false (There is a difference)	Type II Error	Correct conclusion ($p < 0.05$)

Gaddis ML et al. Ann Emerg Med. 1990; 19:1462-8.

Poll Question

Calculate the beta-value necessary to achieve a power of 87%

Poll Question

In the results of a randomized, double-blind, controlled clinical trial, the difference in hospital readmission rates between the intervention group and the control group is 8% ($p=0.02$), and it is concluded that there is a statistically significant difference between the groups.

See next slide for question

Poll Question

Which statement is most consistent with this finding and conclusions?

- A. The chance of making a type I error is 2 in 100
- B. There is a high likelihood of having made a type II error
- C. The trial does not have enough power
- D. The chance of making a type I error is 5 in 100

Confidence Intervals

- A confidence interval is essentially a range of values that you believe the “true” value lies between
 - Often seen as:
 - (95% CI 8% - 42%)
 - (0.95 CI 0.08, 0.42)
 - Traditionally set at 95%, but can be set at different values as well
 - **Confidence Interval** = $1 - \alpha$ (**Type I Error**)
 - If $\alpha = 0.05 \rightarrow 95\% \text{ CI}$
 - If $\alpha = 0.01 \rightarrow 99\% \text{ CI}$



Confidence Intervals – How do you interpret them?

Purpose of a CI is to determine if significance has been reached

✓ Comparing **DIFFERENCE** data

- Means
 - Result is statistically significant IF the CI range does not ENCOMPASS zero
 - (0.0001 – 2.382) = Statistically significant
 - (-0.221 – 0.31) = NOT statistically significant

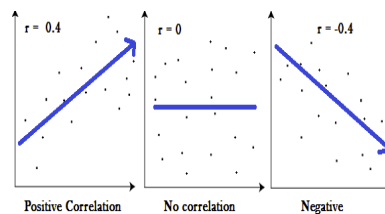
✓ Comparing **RATIO** Data

- RR – Relative Risk
- OR – Odds Ratio
- HR – Hazard Ratio
 - Result is statistically significant IF the CI does not ENCOMPASS one
 - (0.54 – 1.33) = NOT statistically significant
 - (0.22 – 0.99) = Statistically significant



Correlation - Relationship Between Two or More Variables

- Pearson Correlation Coefficient (r)
 - Values can vary from -1 to 1
 - **If the coefficient = 0**
 - The two variables have no relationship or correlation
 - **If coefficient is +**
 - The two variables tend to increase or decrease together
 - **If the coefficient is -**
 - The two variables are inversely related
 - **If the coefficient is 1 or -1**
 - The two variables have a perfect correlation forming a straight line



<https://www.statisticshowto.datasciencecentral.com/probability-and-statistics/correlation-analysis/>



Poll Question

Which of the following is true in regards to a 95% confidence interval?

- If $p < 0.05$ it is considered statistically significant
- Alpha is set at 0.01
- Beta is set at 0.05



Poll Question

Which of the following results is NOT statistically significant?

Glycated hemoglobin in patients with diabetes — %	-0.28±0.03	—	-0.12±0.03	—	-0.16 (-0.25 to -0.08)	A
Hematocrit (%)	1.98±0.10	—	-0.38±0.10	—	2.36 (2.08 to 2.63)	B
Median NT-proBNP (IQR) — pg/mL	-244 (-890 to 260)	—	-141 (-784 to 585)	—	0.87 (0.82 to 0.93)	C
Body weight — kg	-0.73±0.13	—	0.08±0.13	—	-0.82 (-1.18 to -0.45)	D
Systolic blood pressure — mm Hg	-2.4±0.4	—	-1.7±0.4	—	-0.7 (-1.8 to 0.4)	E

Question on next slide

Poll Question

From the previous chart, which of the following results is NOT statistically significant?

- A. Glycated hemoglobin in patients with diabetes
- B. Hematocrit
- C. Median NT-proBNP
- D. Body weight
- E. Systolic blood pressure

RISK

• Risk

- Probability of event or how likely an event is to occur when an intervention is administered

$$\text{Risk} = \frac{\text{\# of subjects with unfavorable event in that arm}}{\text{TOTAL \# of subjects in that arm}}$$

	Outcome Yes	Outcome No
Intervention (Y)	a	b
Control (X)	c	d

$$\text{Risk in Y} = Y = a/(a + b)$$

$$\text{Risk in X} = X = c/(c + d)$$

RISK

• Relative Risk (RR)

- Ratio of the incidence of a given disease in exposed or at-risk persons to the incidence of disease in unexposed persons

$$RR = \frac{\text{Risk of outcome in experimental group}}{\text{Risk of outcome in control group}} = \frac{Y}{X}$$

- Measures the strength of the association between an exposure or intervention and an outcome

• Relative Risk Reduction (RRR) or Increase (RII)

- Reduction in risk with a new therapy relative to the risk without new therapy

$$RRR = (1 - RR) \times 100$$

$$RII = (RR - 1) \times 100$$

	Outcome Yes	Outcome No
Intervention (Y)	a	b
Control (X)	c	d

$$\text{Risk in Y} = Y = a/(a + b)$$

$$\text{Risk in X} = X = c/(c + d)$$

RISK

	Outcome Yes	Outcome No
Intervention (Y)	a	b
Control (X)	c	d

Risk in Y = $a/(a+b)$
Risk in X = $c/(c+d)$

• Absolute Risk Reduction (ARR) or Increase (ARI)

- Attributable Risk

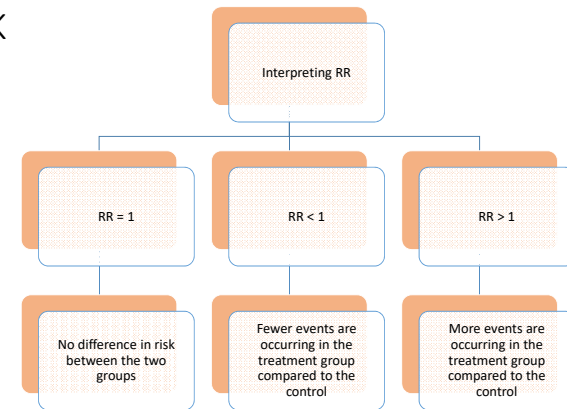
- $ARR = (\% \text{ risk in control group}) - (\% \text{ risk in treatment group})$

- $ARR = X - Y$

- $ARI = (\% \text{ risk in treatment group}) - (\% \text{ risk in control group})$

- $ARI = Y - X$

RISK



Odds

• Odds Ratio (OR)

- Prevalence of EXPOSURE in group with an outcome versus group without an outcome

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

	Outcome Yes	Outcome No
Yes (Drug Group)	a	b
No (Placebo)	c	d

Interpretation of RR/OR

RR	OR	Interpretation
0.75	0.75	25% reduction in risk/odds
1.0	1.0	No difference in risk/odds
1.5	1.5	50% increase in risk/odds
3.0	3.0	3-fold or 200% increase in risk/odds

Poll Question

A randomized controlled trial compares drospirenone and levonorgestrel for risk of thromboembolism (VTE). Authors found a VTE RR 1.7 (95% CI, 1.31-2.19) with drospirenone compared to levonorgestrel. Which one of the following statements is correct?

- A. There is no statistical significance
- B. The risk of VTE is higher with drospirenone
- C. The risk of VTE is higher with levonorgestrel
- D. There is a 30% risk of VTE with levonorgestrel

Hazard Ratio (HR) and Its Interpretation

• Hazard Ratio (HR)

$$HR = \frac{\text{Hazard rate in treatment group}}{\text{Hazard rate in control group}}$$

• Interpreting HR

- **HR = 1**
 - Same event rates in both arms OVER time
- **HR < 1**
 - At any given time, relatively FEWER patients in the treatment group have had an event compared to the control group
- **HR > 1**
 - At any given time, relatively MORE patients in the treatment group have had an event compared to the control group

NNT and NNH

- Calculated when there are significant results reported
- **NNT**
 - Number of patients needed to be treated with intervention in order for ONE patient to achieve a positive result
- **NNH**
 - Number of patients needed to be treated with intervention in order for ONE patient to be affected by an adverse effect

$$NNT = \frac{1}{ARR} \quad [\text{Round } \uparrow]$$

$$NNH = \frac{1}{ARI} \quad [\text{Round } \downarrow]$$

ARR and ARI must be expressed as decimals OR multiply by 100

Poll Question

METHODS

We randomly assigned 3297 patients with type 2 diabetes who were on a standard-care regimen to receive once-weekly semaglutide (0.5 mg or 1.0 mg) or placebo for 104 weeks. The primary composite outcome was the first occurrence of cardiovascular death, nonfatal myocardial infarction, or nonfatal stroke. We hypothesized that semaglutide would be noninferior to placebo for the primary outcome. The noninferiority margin was 1.8 for the upper boundary of the 95% confidence interval of the hazard ratio.

See next slide for question

Poll Question

What was the primary outcome of the SUSTAIN-6 trial?

Poll Question

Looking at this table from the SUSTAIN-6 trial, calculate the hazard ratio for the primary outcome

Outcome	Semaglutide (N=1648)		Placebo (N=1649)	
	no. (%)	no./100 person-yr	no. (%)	no./100 person-yr
Primary composite outcome†	108 (6.6)	3.24	146 (8.9)	4.44
Expanded composite outcome‡	199 (12.1)	6.17	264 (16.0)	8.36
All-cause death, nonfatal myocardial infarction, or nonfatal stroke	122 (7.4)	3.66	158 (9.6)	4.81
Death				
From any cause	62 (3.8)	1.82	60 (3.6)	1.76
From cardiovascular cause	44 (2.7)	1.29	46 (2.8)	1.35
Nonfatal myocardial infarction	47 (2.9)	1.40	64 (3.9)	1.92
Nonfatal stroke	27 (1.6)	0.80	44 (2.7)	1.31

See next slide for question

Poll Question

Looking at this table on the previous slide from the SUSTAIN-6 trial, calculate the hazard ratio for the primary outcome

Poll Question

Looking at this table from the SUSTAIN-6 trial, calculate the number needed to treat for the primary outcome

Outcome	Semaglutide (N=1648)		Placebo (N=1649)		Hazard Ratio (95% CI)*
	no. (%)	no./100 person-yr	no. (%)	no./100 person-yr	
Primary composite outcome†	108 (6.6)	3.24	146 (8.9)	4.44	0.74 (0.58-0.95)
Expanded composite outcome‡	199 (12.1)	6.17	264 (16.0)	8.36	0.74 (0.62-0.89)
All-cause death, nonfatal myocardial infarction, or nonfatal stroke	122 (7.4)	3.66	158 (9.6)	4.81	0.77 (0.61-0.97)
Death					
From any cause	62 (3.8)	1.82	60 (3.6)	1.76	1.05 (0.74-1.50)
From cardiovascular cause	44 (2.7)	1.29	46 (2.8)	1.35	0.98 (0.65-1.48)
Nonfatal myocardial infarction	47 (2.9)	1.40	64 (3.9)	1.92	0.74 (0.51-1.08)
Nonfatal stroke	27 (1.6)	0.80	44 (2.7)	1.31	0.61 (0.38-0.99)

Go to next slide for question

Looking at the table on the previous slide from the SUSTAIN-6 trial, calculate the number needed to treat for the primary outcome

Take Away Points

- Identify the types of variables/data presented in a study
 - Nominal, ordinal, continuous
- Distribution of the data
 - Normal versus non-normal, parametric versus nonparametric
- Number of groups
 - 2 groups, before/after, 3 or more groups

Take Away Points

KNOW THIS TABLE

	Independent		Dependent	
Number of groups	2 groups	≥2 groups	2 groups (either crossover or pre/post treatment)	Multiple measures in ≥2 groups
Continuous Data	t-test	ANOVA	Paired t-test	Repeated Measures ANOVA
Ordinal Data	Mann-Whitney U Test (Wilcoxon Rank Sum)	Kruskal-Wallis	Wilcoxon Signed Rank Test	Friedman
Nominal Data	Chi-square or Fisher's Exact Test	Chi-square or Fisher's Exact Test	McNemar	Cochrane Q

Poll Question

On a scale of 1-5, how would you rate your skills in interpreting medical literature?

- 1 - Poor
- 2 - Fair
- 3 - Average
- 4 - Good
- 5 - Excellent

Thank you for coming!
Questions?