

Trial Design and Biostatistics

Katie J. Suda, Pharm.D., M.S.

Research Health Scientist and Associate Director of Clinical Therapeutics
VA Center for Health Equity Research and Promotion

Professor of Medicine and Pharmacy and Therapeutics
Director, Transition to Independence Program
University of Pittsburgh Schools of Medicine and Pharmacy
Pittsburgh, Pennsylvania



1

Conflict of Interest Disclosure

- I have no actual or potential conflicts of interest in relation to this presentation.
- I have or am currently receiving research funding from AHRQ, NIH, VA HSR&D, VA RR&D, and VA QUERI program.
- *Opinions expressed today are those of the presenter and do not represent positions or views of the Department of Veterans Affairs or the U.S. Government.*



2

Learning Objectives

- After completing the lecture, the audience will be able to:
 1. Describe hypothesis testing and state the meaning of and distinguish between p values and confidence intervals and measures of central tendency and data spread.
 2. Define, compare and contrast the concepts of internal and external validity, causation, association, bias, and confounding in trial design. Select strategies to eliminate or control for covariates that may impact study conclusions.
 3. Compare and contrast the advantages and disadvantages of various study designs (e.g., prospective, retrospective, case-control, cohort, cross-sectional, randomized controlled clinical trials, systematic review, meta-analysis)
 4. Determine why a statistical test is appropriate or not appropriate, given the sample distribution, data type, and study design. Interpret statistical and clinical significance for results from commonly used statistical tests.
 5. Define and evaluate odds ratio, risk/incidence rate, relative risk (RR), number needed to treat, number needed to harm, and other risk estimates.

3

Agenda

- Literature evaluation
- Example study
- Study design - RCTs
- Statistics
- Epidemiologic measurements
- Other study designs

4

Study Design and Biostatistics

- Assume you have some baseline understanding of trial design and biostatistics
- Insufficient time to capture everything
 - Focus on the most confusing topics
- Integrate some of the material in the chapter in the context of evaluating a clinical trial
 - No patient cases
 - There is no perfect study

5

30 Evaluative Questions To Ask of a Trial

Acknowledgement:
Dr. Lee Vermeulen
University of Kentucky
UK Healthcare

6

30 Evaluative Questions

- Included in your handout
- Systematic approach to clinical trial evaluation
- Not all questions necessarily of equal weight
- Many questions overlap
- Use as tool to find important flaws in clinical research
- No perfect trial
- Goal: become an informed consumer and use high quality evidence

7

Background

- Fake news
- Flawed research
- Exaggerated study results

8

EXAMPLE STUDY:

Randomized, double-blind, placebo-control study on decolonization procedures for methicillin resistant *Staphylococcus aureus* (MRSA) among HIV-infected adults.

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. PLoS One. 2015;10:e0128071.

9

Was the study design appropriate to the hypothesis?

10

Study Objective and Hypothesis

- “Primary objective was to evaluate the effectiveness of decolonization procedure with nasal mupirocin & hexachlorophene soap compared to placebo among HIV-infected persons on MRSA colonization at 6 months post-randomization.”
- Did not explicitly state the primary hypothesis:
 - Primary H1: There is a difference in MRSA colonization between the decolonization procedure and placebo.
 - H1=Alternative hypothesis (also noted as Ha)
 - Primary H0: There is no difference in MRSA colonization.
 - H0=Null hypothesis

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant Staphylococcus aureus (MRSA) among HIV infected adults. PLoS One. 2015;10:e0128071.

11





Study Objective and Hypothesis

- Secondary objectives were to assess all randomization time points (when MRSA positive), time to MRSA clearance, and the effectiveness of decolonization procedures on subsequent MRSA infections and SSTI.
- Did not explicitly state the secondary hypothesis:
 - Secondary H1: There is a difference in subsequent MRSA infections between the decolonization procedure and placebo.
 - Secondary H0: There is no difference in subsequent MRSA infections.

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant Staphylococcus aureus (MRSA) among HIV infected adults. PLoS One. 2015;10:e0128071.

12

Null vs. Alternative Hypothesis and Decision Errors

Test (study) Result	Underlying Truth or Reality	
	H0 is true (no difference)	H0 is false (difference)
Accept H0 (no difference)	No error (correct decision) 	Type II error (β) 
Reject H0 (difference)	Type I error (α) 	No error (correct decision) 

13

Null vs. Alternative Hypothesis and Decision Errors – Type I Error

Test Result	Underlying Truth or Reality	
	H0 is true (no difference)	H0 is false (difference)
Accept H0 (no difference)	No error (correct decision)	Type II error (β)
Reject H0 (difference)	Type I error (α)	No error (correct decision)

- **Type I error** is to reject the H0 when it is true - “False” alarm
- Likelihood of making a **Type I error** is defined as the significance level (α), typically set at 0.05
 - Reject H0 if the p value is $< \alpha$

14

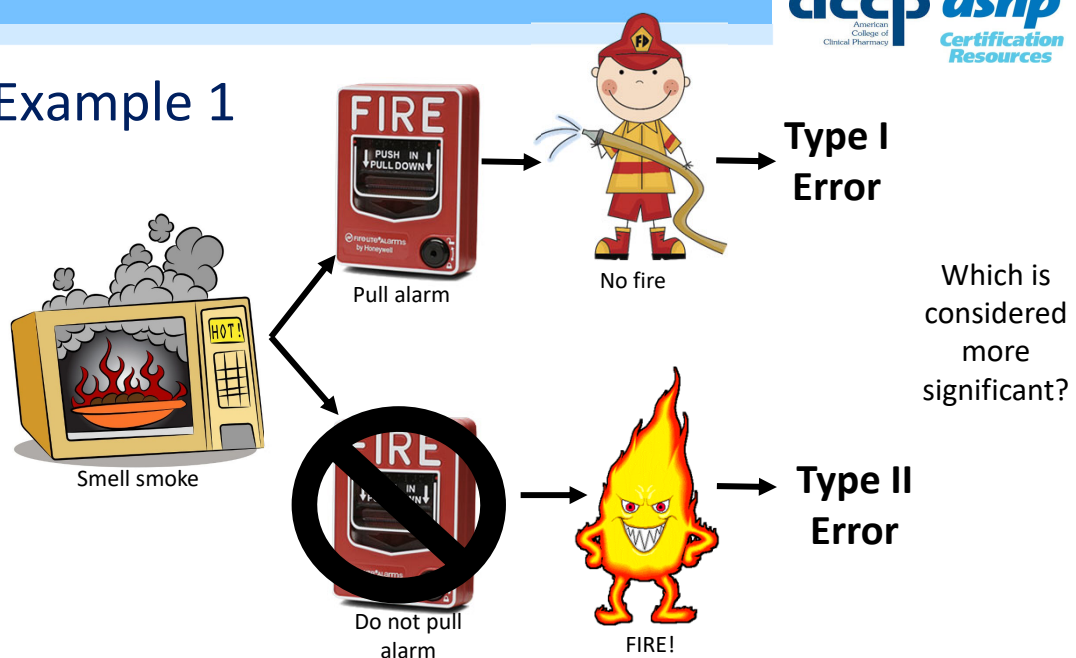
Null vs. Alternative Hypothesis and Decision Errors – Type II Error

Test Result	Underlying Truth or Reality	
	H0 is true (no difference)	H0 is false (difference)
Accept H0 (no difference)	No error (correct decision)	Type II error (β)
Reject H0 (difference)	Type I error (α)	No error (correct decision)

- **Type II error** is to fail to reject (or accept) the H0 when it is false
- Likelihood of making a **Type II error** is defined as beta, typically set at 0.20
 - 1-beta=Power; the ability to detect a difference if one truly exists

15

Example 1



Images adapted under the Creative Commons license.

16

Example 2

Truth	Disease present	Medical Test	Error	Interpretation
Patient <u>does not have</u> influenza	No	Influenza test is <u>positive</u> for disease	Type I Error	False positive
Patient truly <u>has</u> influenza	Yes	Influenza test is <u>negative</u> for disease	Type II Error	False negative

17

Question 1

- In the results of a randomized, double-blind, controlled clinical trial, the difference in 10-year mortality between the intervention group and the control group is 6% ($p = 0.01$), and it is concluded that there is a statistically significant difference between the groups. Which statement is most consistent with this finding and conclusions?
 - The chance of making a type I error is 5 in 100.
 - The trial does not have enough power.
 - There is a high likelihood of having made a type II error.
 - The chance of making a type I error is 1 in 100.

Learning Objective 1
 Domain 3 Task 2

18

Question 1 - Answer

- In the results of a randomized, double-blind, controlled clinical trial, the difference in 10-year mortality between the intervention group and the control group is 6% ($p = 0.01$), and it is concluded that there is a statistically significant difference between the groups. Which statement is most consistent with this finding and conclusions?
 - A. The chance of making a type I error is 5 in 100.
 - B. The trial does not have enough power.
 - C. There is a high likelihood of having made a type II error.
 - D. The chance of making a type I error is 1 in 100.**

Learning Objective 1
 Domain 3 Task 2

19

Question 1 – Answer Explained

- In the results of a randomized, double-blind, controlled clinical trial, the difference in 10-year mortality between the intervention group and the control group is 6% ($p = 0.01$), and it is concluded that there is a statistically significant difference between the groups. Which statement is most consistent with this finding and conclusions?
 - A. The chance of making a type I error is 5 in 100.
 - B. The trial does not have enough power.
 - C. There is a high likelihood of having made a type II error.
 - D. The chance of making a type I error is 1 in 100.**

Answer A is incorrect. The typical a priori error (type I) rate is 5% (i.e., when the study was designed, the error rate was designed to be 5% or less). The actual type I error rate is reported in the question as 0.01 (1%).

Answer B is incorrect. Information is insufficient to select Answer B.

Answer C is incorrect. A type II error was not made because this error has to do with not finding a difference when one truly exists.

Answer D is correct. In this question, the type I error rate is 1%, the value of the p value.

Learning Objective 1
 Domain 3 Task 2

20

Null vs. Alternative Hypothesis and Decision Errors

- Primary H1: There is a difference in MRSA colonization between the decolonization procedure and placebo.
- Alternative hypothesis (H1): States that there is a difference between groups being compared (treatment A does not equal treatment B)

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

21

Null vs. Alternative Hypothesis and Decision Errors

- Primary H0: There is no difference in MRSA colonization.
 - Null hypothesis (H0): No difference between groups being compared
 - ‘Fail to reject’ H0 (or accepting H0) means that no difference exists between groups.
 - Reject H0: statistically significant difference between groups (unlikely attributable to chance)
 - Fail to reject H0: no statistically significant difference between groups (any apparent differences may be attributable to chance). We are not concluding that the treatments are equal.

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

22

Was the study design appropriate to the hypothesis?

- “....A randomized, double-blinded, placebo-controlled clinical study was designed....” (Pg 6)
- To determine a cause-and-effect relationship between treatments, an RCT is appropriate.

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

23

Study design – RCT

- Randomized controlled clinical trial (RCT)
- Prospective, experimental, hypothesis testing
- Patients are randomly assigned to:
 - Intervention – “experimental” treatment
 - Control – “nonexperimental” treatment
- Most reliable results that are likely to impact patient care
- GOLD STANDARD for studying interventions

24

Study design – Parallel RCT

- Parallel designed RCT
- Majority of RCT
- Separate patients into two or more groups
 - One group control, the other receives the active intervention

Nasal mupirocin + soap (intervention)

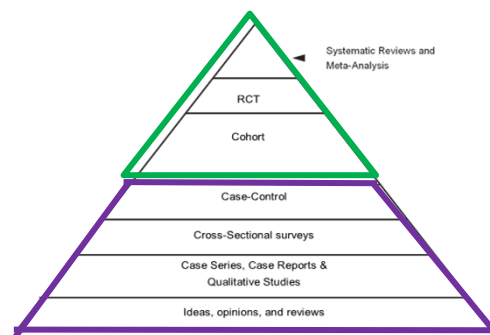
Placebo (control)

END (and/or randomized again if positive)

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

25

Study design – Hierarchy of evidence



Determine
cause and
effect
relationships
(causal)

Determine
associations

26

Was the study design appropriate to the hypothesis?



27

Was the number of subjects enrolled adequate?



28

Sample Size

Statistical Methods

The sample size for this study was based on the assumption that 10% of the HIV-infected persons would be colonized during the study period [3,31,32] and that 85% in the placebo arm and 39% of the treatment arm would remain colonized based on prior literature [22], hence the sample size was estimated as 420 participants with a power 80% and alpha level of 0.05. Given potential loss to follow-up and that the colonization and clearance rates may vary, the study enrollment was *a priori* set at 550 participants. The sample size was also deemed adequate for the MRSA infection/SSTI outcome assuming 38% of the placebo group and 10% of the treatment group would develop infection based on a prior military study [13] and a subsequent study in HIV-infected persons [15], and accounting for the 6-month study visits over a two-year period.

Image adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

29

Sample Size

- To calculate sample size:
 - Study design (type, matching)
 - Decisional thresholds for alpha and beta
 - Anticipated variability for the primary outcome
 - Effect size
 - Anticipated size of the difference in the primary outcome between groups
 - Attrition (drop out)

30

Sample Size

- To calculate sample size:
 - Study design (type, matching); RCT (parallel)
 - Decisional thresholds for alpha and beta; $\alpha=0.05$; $\beta=0.2$
 - Anticipated variability for the primary outcome; ???
 - May have used variability in prevalence between studies (3 were cited)
 - Effect size; $85\% - 39\% = 46\%$
 - Attrition (drop out); 30% dropout

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

31

Sample Size

Statistical Methods

The sample size for this study was based on the assumption that 10% of the HIV-infected persons would be colonized during the study period [3,31,32] and that 85% in the placebo arm and 39% of the treatment arm would remain colonized based on prior literature [22], hence the sample size was estimated as 420 participants with a power 80% and alpha level of 0.05. Given potential loss to follow-up and that the colonization and clearance rates may vary, the study enrollment was *a priori* set at 550 participants. The sample size was also deemed adequate for the MRSA infection/SSTI outcome assuming 38% of the placebo group and 10% of the treatment group would develop infection based on a prior military study [13] and a subsequent study in HIV-infected persons [15], and accounting for the 6-month study visits over a two-year period.

30% drop out

Secondary outcome - accounts for study design (repeated measures)

Primary Outcome Effect Size =
 $85\% - 39\% = 46\%$

$\alpha=0.05$

$\beta=0.2$

Image adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

32

Was the number of subjects enrolled adequate?



33



Question 2

- An investigational drug is being compared with an existing drug for the treatment of diabetes in subjects ≥ 65 years of age. The study is designed to detect a minimum 20% difference in response rates between the groups, if one exists, with an a priori alpha (α) of 0.05. The investigators believe that a smaller difference (10% difference vs. 20% difference) would be more clinically meaningful. In revising their study, they decide they want to be able to detect a minimum 10% difference in response. Which change to the study parameters is most appropriate?
- A. Increase the sample size.
 - B. Change the study design to compare the investigational drug with placebo.
 - C. Select an alpha (α) of 0.01 as a cutoff for statistical significance.
 - D. Decrease the sample size.

Learning Objective 1
Domain 3 Task 2

34

Question 2 – Answer

- An investigational drug is being compared with an existing drug for the treatment of diabetes in subjects ≥ 65 years of age. The study is designed to detect a minimum 20% difference in response rates between the groups, if one exists, with an a priori alpha (α) of 0.05. The investigators believe that a smaller difference (10% difference vs. 20% difference) would be more clinically meaningful. In revising their study, they decide they want to be able to detect a minimum 10% difference in response. Which change to the study parameters is most appropriate?
- Increase the sample size.**
 - Change the study design to compare the investigational drug with placebo.
 - Select an alpha (α) of 0.01 as a cutoff for statistical significance.
 - Decrease the sample size.

Learning Objective 1
 Domain 3 Task 2

35

Question 2 – Answer Explained

- An investigational drug is being compared with an existing drug for the treatment of diabetes in subjects ≥ 65 years of age. The study is designed to detect a minimum 20% difference in response rates between the groups, if one exists, with an a priori alpha (α) of 0.05. The investigators believe that a smaller difference (10% difference vs. 20% difference) would be more clinically meaningful. In revising their study, they decide they want to be able to detect a minimum 10% difference in response. Which change to the study parameters is most appropriate?

- Increase the sample size.**
- Change the study design to compare the investigational drug with placebo.
- Select an alpha (α) of 0.01 as a cutoff for statistical significance.
- Decrease the sample size.

Answer A is correct. Detecting the smaller difference between the treatments requires more power. Power can be increased in several different ways. Answer A is correct because the most common approach is to increase the sample size.

Answer B is incorrect. Although the subjects are older patients with diabetes, Answer B is likely unethical because treatment options for diabetes are available and thus should not be withheld from study subjects.

Answer C is incorrect. Power can also be increased by increasing the α value, but doing so increases the chances of a type I error. Answer C decreases α , thus making it more difficult to detect differences between groups.

Answer D is incorrect because smaller sample sizes diminish a study's ability to detect differences between groups.

Learning Objective 1
 Domain 3 Task 2

36

Was allocation to treatment groups truly random?



37

Randomization



(Pg 4)

Participants with a positive culture for MRSA colonization at any of the sampled body sites were randomized in a 1:1 fashion to mupirocin (Bactroban) nasal ointment plus hexachlorophene (pHisoHex) soap for seven days or to a placebo nasal ointment and body soap for seven days. Study procedures, including randomization, are shown in [Fig 1](#). The study was double-blinded with neither the participant nor physician being aware of the treatment assignment, and randomization occurred in blocks of ten at each clinical site. Randomized subjects received

Image adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

38

Randomization

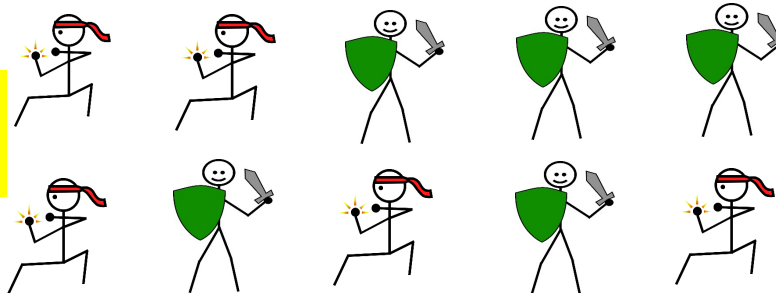
- What is randomization?
 - Each subject has an equal and independent chance of being allocated into either group
 - Attempts to balance observed and unobserved variables
- Block randomization
- Stratified randomization
- Cluster randomization

39

Block Randomization

- Study assignment is based on a sequence of study treatment/control (called 'blocks') to obtain equal sample sizes in each group
- Blocks of 10 = TTTTCCCCTC

T=Participant assigned to treatment
C=Assigned to the control group

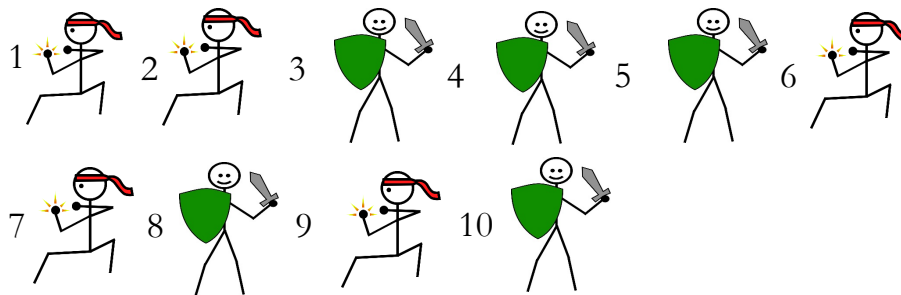


Images adapted under the Creative Commons license.

40

Block Randomization

- Study assignment is based on a sequence of study treatment/control (called 'blocks') to obtain equal sample sizes in each group
- Block = TTTTCCCCTC

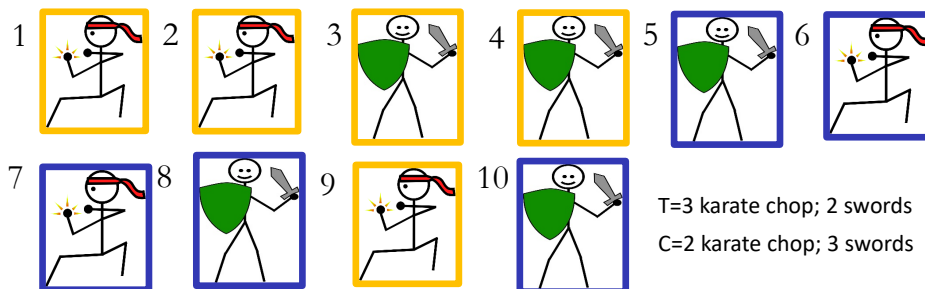


Images adapted under the Creative Commons license.

41

Block Randomization

- Study assignment is based on a sequence of study treatment/control (called 'blocks') to obtain equal sample sizes in each group
- Block = TTTTCCCCTC



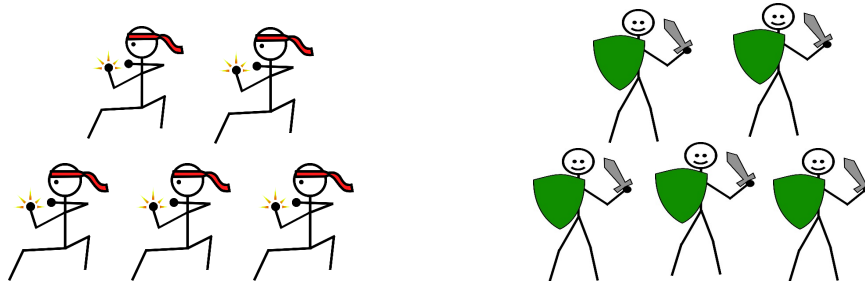
T=3 karate chop; 2 swords
C=2 karate chop; 3 swords

Images adapted under the Creative Commons license.

42

Stratified Randomization

- Group participants by clinical feature(s) that may influence outcome (“strata”); randomization occurs within each stratum.

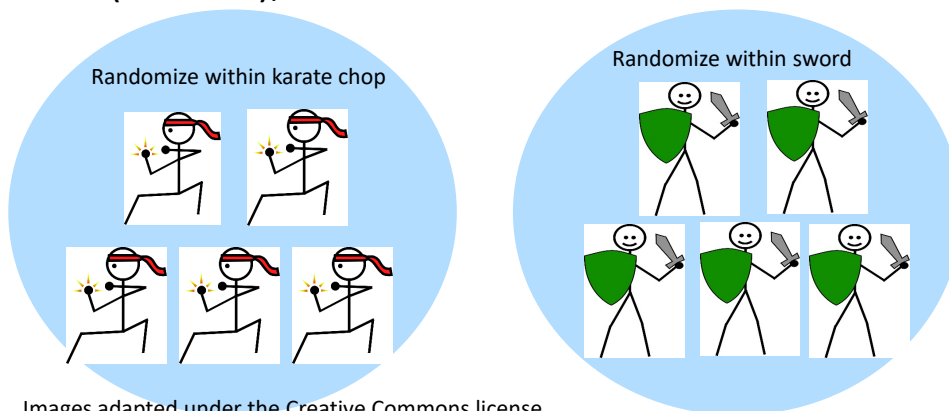


Images adapted under the Creative Commons license.

43

Stratified Randomization

- Group participants by clinical feature(s) that may influence outcome (“strata”); randomization occurs within each stratum.

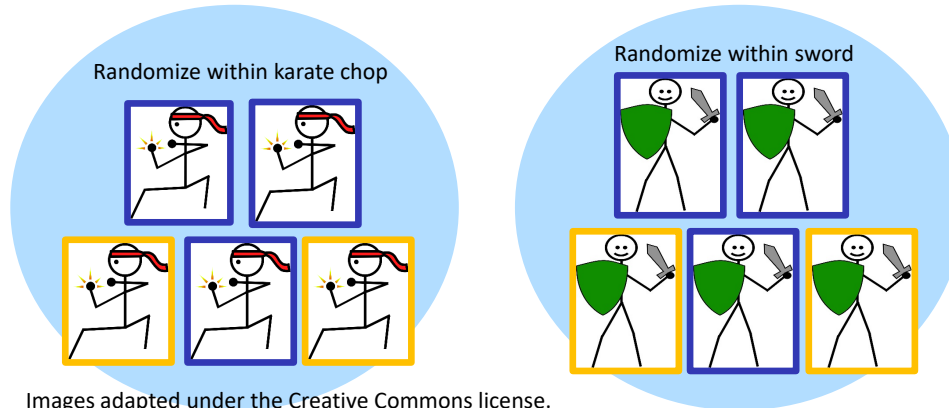


Images adapted under the Creative Commons license.

44

Stratified Randomization

- Group participants by clinical feature(s) that may influence outcome (“strata”); randomization occurs within each stratum.

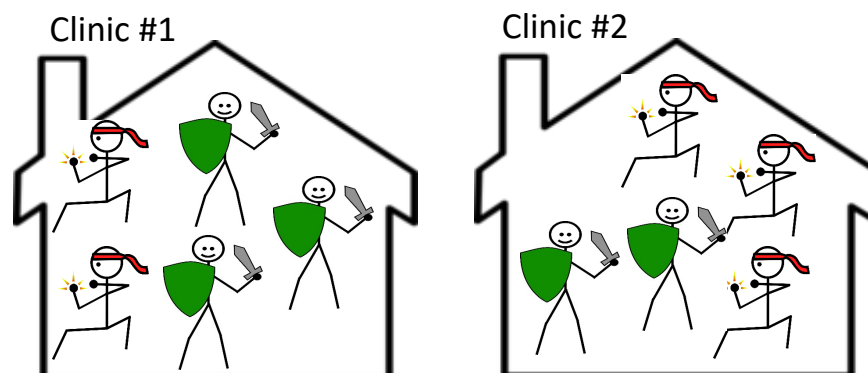


Images adapted under the Creative Commons license.

45

Cluster Randomization

- Groups of subjects are randomized (vs the individual subject being randomized); is commonly done for clinics or hospital

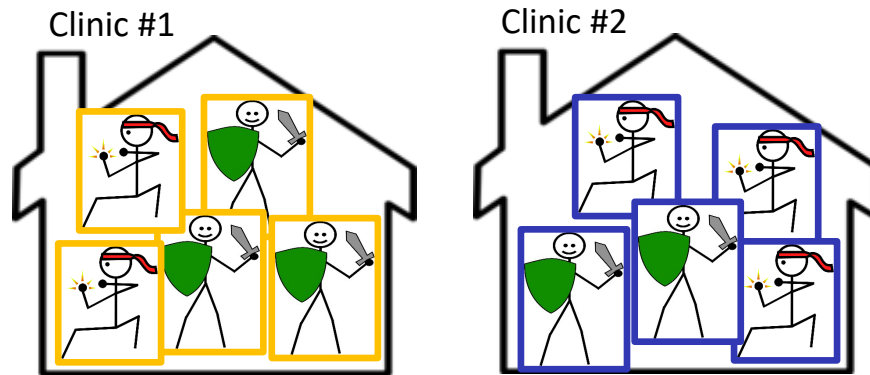


Images adapted under the Creative Commons license.

46

Cluster Randomization

- Groups of subjects are randomized (vs the individual subject being randomized); is commonly done for clinics or hospital



Images adapted under the Creative Commons license.

47

Was allocation to treatment groups truly random?

48

Was the study blinded? Was it truly blinded?



49

Blinding

- “Placebos were of similar appearance, but without specific antibacterial activity.” (Pg 4)

days. Study procedures, including randomization, are shown in [Fig 1](#). The study was double-blinded with neither the participant nor physician being aware of the treatment assignment, and randomization occurred in blocks of ten at each clinical site. Randomized subjects received

Image adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

50

Blinding

- Study states that it is ‘double blinded’
 - Patients and physicians unaware of treatment selection
 - Both subjects and investigators are unaware of subject assignment to active/control.
 - Intervention and placebo identical in appearance and patient instructions
- However, may have been ‘triple blinded’
 - “Investigators remained blinded throughout the study”

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

51

Was the study blinded? Was it truly blinded?

52

Was an appropriate statistical method used?



53

Statistical Analysis Pg 6

Descriptive statistics were performed evaluating the baseline characteristics of the study population and presented as numbers (percentages) and medians (interquartile ranges, IQR) for categorical and continuous variables, respectively. The primary outcome was evaluated using the Fisher's exact test. An analysis which included all randomizations per participant during the 2-year study period (including multiple randomizations) was performed using GEE models to account for subject level correlations. Considering the outcome as a binary structure, logit link function was used; an unstructured type was chosen to test for correlations, which means that no restrictions were placed on the correlations.

Additionally, secondary analyses included time to MRSA clearance using monthly swab data and effectiveness of study drug on subsequent MRSA infection/SSTI, which were evaluated using logrank test and time-to-event Cox proportional hazards models. Finally, univariable and multivariable logistic regressions models were created to evaluate for predictors (including demographics, clinical factors such as CD4 counts, and treatment group) for MRSA clearance at the 6-month time point. All statistical analyses were performed using SAS version 9.2 (SAS Institute, Cary, NC).

Image adapted under the Creative Commons license. Weintrob A, Bebu J, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

54

Statistical Analysis

- Descriptive statistics
 - Visual methods for displaying data
 - Measures of central tendency (e.g, mean)
 - Measures of data spread and variability (e.g, SD)
- Inferential statistics
 - Conclusions or generalizations made about a population (large group) from the study of a sample of that population

55

Statistical Analysis

- To evaluate appropriateness of statistical testing, you need the following:
 - Type of data (levels of measure)
 - Distribution of data
 - Number of groups
 - Study design
 - Was matching used (e.g, cross-over studies, case control)?
 - Number of observations (usually subjects)
 - Presence of confounders

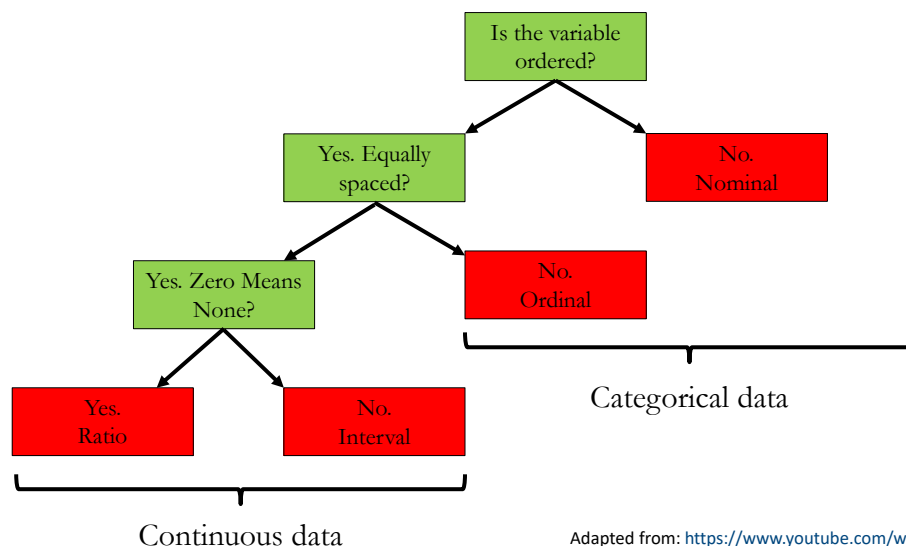
56

Statistics – Levels of measure

- Continuous Data: Number data with a consistent change in magnitude between units
 - Interval: Zero point is arbitrary (e.g., 0° F does not mean “no heat”)
 - Ratio: Has an absolute zero (e.g., heart rate)
- Nominal data: Counts of categories/groups without a rank (e.g., sex)
- Ordinal Data: Has a natural order, but they are an arbitrary and there is not a defined distance between one value and another

57

Statistics – Levels of Measure

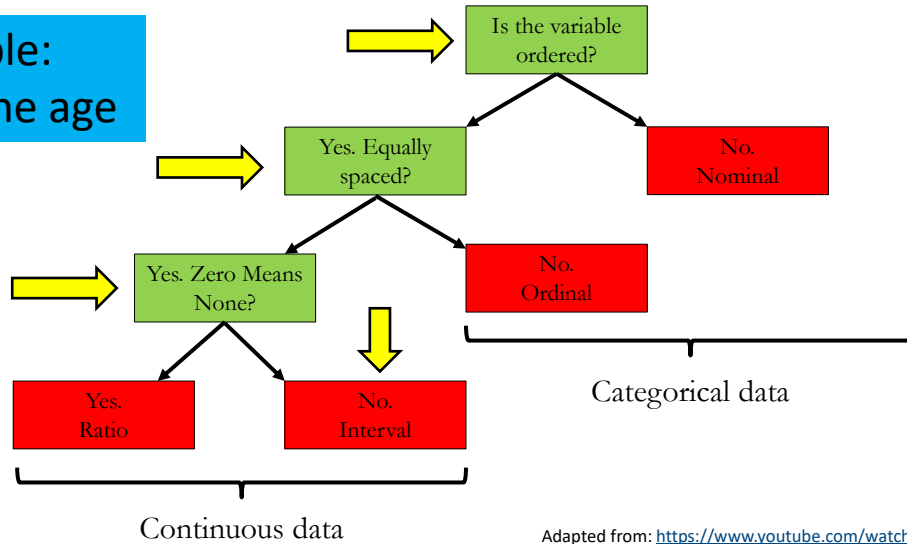


Adapted from: <https://www.youtube.com/watch?v=LPHYXPBK ks>.

58

Statistics – Levels of Measure

Example:
Baseline age

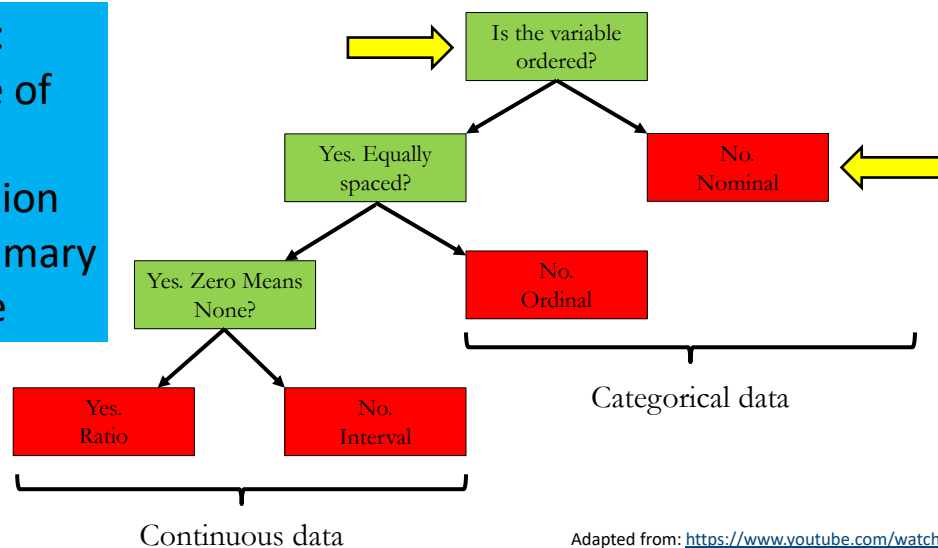


Adapted from: https://www.youtube.com/watch?v=LPHYXPBK_ks.

59

Statistics – Levels of Measure

Example:
Presence of
MRSA
colonization
(Y/N); Primary
Outcome

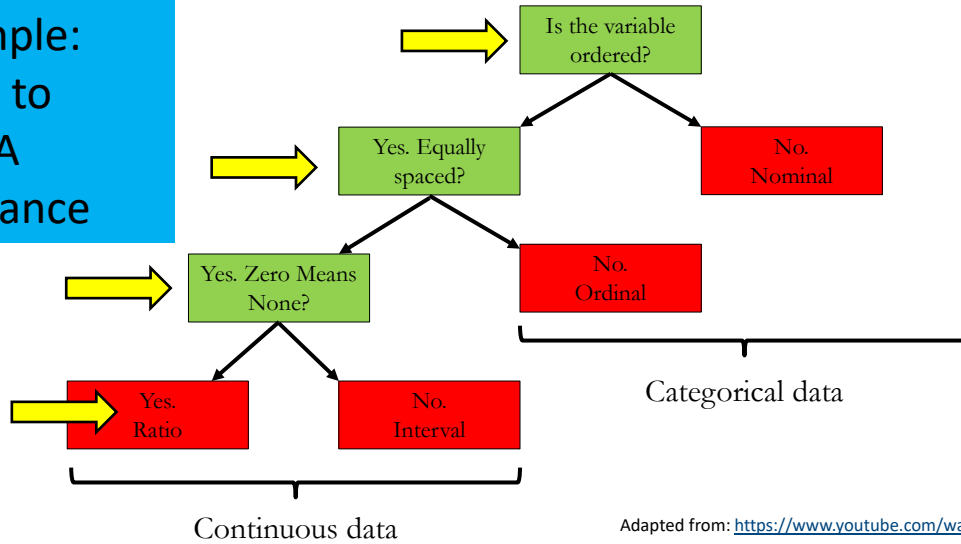


Adapted from: https://www.youtube.com/watch?v=LPHYXPBK_ks.

60

Statistics – Levels of Measure

Example:
Time to
MRSA
clearance



Adapted from: https://www.youtube.com/watch?v=LPHYXBK_ks.

61

Variable	Continuous or Categorical	Levels of Measure
Baseline age	Continuous	Interval
Gender	Categorical	Nominal
CD4 count	Continuous	Ratio
CD4 count by category	Categorical	Nominal
History of MRSA infection	Categorical	Nominal
Presence of MRSA colonization	Categorical	Nominal
Time to MRSA clearance	Continuous	Ratio
Effectiveness of study drug on future colonization (time)	Continuous	Ratio

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

62

Statistical Analysis

- To evaluate appropriateness of statistical testing, you need the following:
 - Type of data (levels of measure)
 - Distribution of data
 - Number of groups
 - Study design
 - Was matching used (e.g, cross-over studies, case control)?
 - Number of observations (usually subjects)
 - Presence of confounders

63

Statistics – Parametric or Non-parametric

- Is the data normally distributed?
- For continuous data
- How to know:
 - View the graph
 - Median = mean
 - Formal statistical test

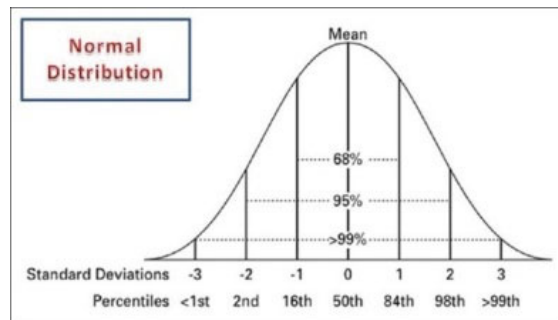


Image adapted under the Creative Commons license: Ali Z, Bhaskar SB. Basic statistical tools in research and data analysis. *Indian J Anesth.* 2016;60:662-669.

64

Statistics – Parametric or Non-parametric

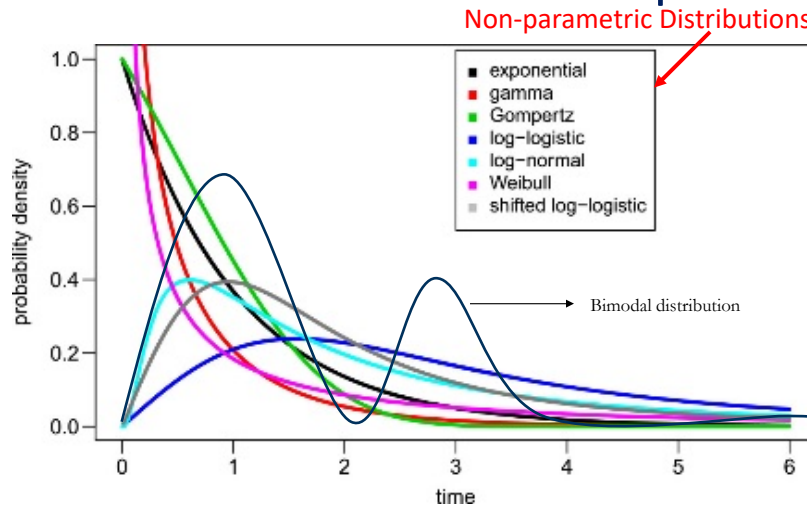


Image adapted under the Creative Commons license: Lover AA, Zhao X, Gao Z, et al. The distribution of incubation and relapse times in experimental human infections with the malaria parasite *P. vivax*. *BMC Infect Dis.* 2014;14:539.

65

Statistics – Parametric or Non-parametric

- Why does this matter?
 - Says that your data reflects the underlying population
 - Assumption for robust (more powerful) parametric statistical tests
 - Parametric tests are generally more interpretable

66

Variable	Continuous or Categorical	Levels of Measure	Normally distributed? (only for continuous data)
Baseline age	Continuous	Interval	Non-parametric
Gender	Categorical	Nominal	---
CD4 count	Continuous	Ratio	Non-parametric
CD4 count by category	Categorical	Nominal	---
History of MRSA infection	Categorical	Nominal	---
Presence of MRSA colonization (1° & 2° outcome)	Categorical	Nominal	---
Time to MRSA clearance	Continuous	Ratio	???
Effectiveness of study drug on future colonization time	Continuous	Ratio	???


Adapted under the Creative Commons license: Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

67


Statistical Analysis

- To evaluate appropriateness of statistical testing, you need the following:
 - Type of data
 - Distribution of data
 - Number of groups
 - Study design
 - Was matching used (e.g, cross-over studies, case control)? Were subjects paired? Were there repeated measurements (repeated measures)? Is it a before-and-after study? Are the groups independent of each other?
 - Number of observations (usually subjects)
 - Presence of confounders

68

Statistics – Bivariate					
	Data Type	2 groups (independent)	3+ groups (independent)	Before and after a single treatment in the same subjects (2 groups)	Multiple treatments in the same subjects (3+groups)
Non-parametric	Nominal	Chi Square; Fisher's exact test (< 5 in any cell or N < 20)	Chi Square; Fisher's exact test (< 5 in any cell or N < 20)	McNemar's test	Cochrane's Q test; Mantel-Haenszel test
	Ordinal or nonparametric continuous, not normally distributed	Mann-Whitney U test; Wilcoxon rank sum test; Wilcoxon Mann-Whitney test	Kruskal-Wallis test	Wilcoxon signed rank test	Friedman ANOVA by ranks
Parametric	Continuous, normally distributed	(Student's) t-test	ANOVA followed by post-hoc testing (e.g., Bonferroni, Turkey)	Paired t-test	Repeated-measures ANOVA


69

Statistics – AGE					
	Data Type	2 groups (independent)	3+ groups (independent)	Before and after a single treatment in the same subjects (2 groups)	Multiple treatments in the same subjects (3+groups)
Non-parametric	Nominal	Chi Square; Fisher's exact test (< 5 in any cell or N < 20)	Chi Square; Fisher's exact test (< 5 in any cell or N < 20)	McNemar's test	Cochrane's Q test; Mantel-Haenszel test
	Ordinal or nonparametric continuous, not normally distributed	Mann-Whitney U test; Wilcoxon rank sum test; Wilcoxon Mann-Whitney test	Kruskal-Wallis test	Wilcoxon signed rank	Friedman ANOVA by ranks
Parametric	Continuous, normally distributed	(Student's) t-test	ANOVA followed by post-hoc testing (e.g., Bonferroni, Turkey)		

Why isn't the correct test a t-test? In this example, age is non-parametric. If age was normally distributed (parametric), a t-test would have been appropriate.


Remember: You always want to use the most robust test. Parametric tests are more robust than non-parametric.

70

					
Variable	Con't or Categorical	Levels of Measure	Normal dis?	Statistical Test	
Baseline age	Continuous	Interval	Non-para	Wilcoxon Mann Whitney	
Gender	Categorical	Nominal	---		
CD4 count	Continuous	Yes	Non-param		
CD4 count by category	Categorical	Nominal	---		
History of MRSA infection	Categorical	Nominal	---		
Presence of MRSA colonization (1° & 2° outcome)	Categorical	Nominal	---		
Time to MRSA clearance	Continuous	Ratio	???		
Effectiveness of study drug on future colonization time	Continuous	Ratio	???		

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

71

Statistics – GENDER					
	Data Type	2 groups (independent)	3+ groups (independent)	Before and after a single treatment in the same subjects (2 groups)	Multiple treatments in the same subjects (3+groups)
Non-parametric	Nominal	Chi Square; Fisher's exact test (< 5 in any cell or N < 20)	Chi Square; Fisher's exact test (< 5 in any cell or N < 20)	McNemar's test	Cochrane's Q test; Mantel-Haenszel test
	Ordinal or nonparametric continuous, not normally distributed	Mann-Whitney U test; Wilcoxon rank sum test; Wilcoxon Mann-Whitney test	Kruskal-Wallis test	Wilcoxon	Friedman
Parametric	Continuous, normally distributed	(Student's) t-test	ANOVA followed by post-hoc testing (e.g., Bonferroni, Turkey)	Paired t-test	Repeated-measures ANOVA

Matched or paired

Why isn't the correct test a Chi Square?
In this example, there were <5 women in the study group. If there were ≥5 in every cell and a total N >20, a chi square test would have been appropriate.

72

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

Variable	Con't or Categorical	Levels of Measure	Normal dis?	Statistical Test
Baseline age	Continuous	Interval	Non-para	Wilcoxon Mann Whitney
Gender	Categorical	Nominal	---	Fisher's exact test
CD4 count	Continuous	Ratio	Non-param	Wilcoxon Mann Whitney
CD4 count by category	Categorical	Nominal	---	Chi Square
History of MRSA infection	Categorical	Nominal	---	Chi Square
Presence of MRSA colonization (outcome)	Categorical	Nominal	---	Fisher's exact test (primary); GEE (secondary)
Time to MRSA clearance	Continuous	Ratio	???	Logrank; Cox PH
Effectiveness of study drug on future colonization time	Continuous	Ratio	???	Logrank; Cox PH

73

Statistical Analysis

- To evaluate appropriateness of statistical testing, you need the following:
 - Type of data
 - Distribution of data
 - Number of groups
 - Study design
 - Was matching used (e.g, cross-over studies, case control)?
 - Number of observations (usually subjects)
 - Presence of confounders
 - History of MRSA colonization

74

Question 3

- A single group of patients are given a new medication for Hepatitis C. The effects of the medication are evaluated by assessing viral loads in the same patient before and after the addition of the new medication. The outcome of interest is normally distributed and has a nearly identical variance before and after the addition of the new drug. Which of the following is the best statistical inference test for comparing the efficacy of the new medication on viral load in the study?
- A. One-way ANOVA
 - B. Mantel-Haenszel test
 - C. Paired t-test
 - D. Independent t-test

Learning Objective 6
Domain 3 Task 2

75

Question 3 – Answer

- A single group of patients are given a new medication for Hepatitis C. The effects of the medication are evaluated by assessing viral loads in the same patient before and after the addition of the new medication. The outcome of interest is normally distributed and has a nearly identical variance before and after the addition of the new drug. Which of the following is the best statistical inference test for comparing the efficacy of the new medication on viral load in the study?
- A. One-way ANOVA
 - B. Mantel-Haenszel test
 - C. Paired t-test**
 - D. Independent t-test

Learning Objective 6
Domain 3 Task 2

76

Question 3 – Answer Explained

- A single group of patients are given a new medication for Hepatitis C. The effects of the medication are evaluated on the patient before and after the addition of the new drug. The variable of interest is normally distributed and homogeneous. Which statistical inference test for comparing two groups on viral load in the study?

- A. One-way ANOVA
- B. Mantel-Haenszel test
- C. Paired t-test**
- D. Independent t-test

Answer A is incorrect: Viral load is a continuous parametric variable. The One-way ANOVA test is best used for 3 or more independent groups.

Answer B is incorrect: Mantel-Haenszel tests are for nominal data.

Answer C is correct: T-tests are used for continuous data; the paired t-test is for two dependent (or matched) groups

Answer D is incorrect: The independent t-test is used for continuous independent groups. The observations are considered dependent because the outcome (viral load) is measure before and after administration of the study drug on the same subject.

Learning Objective 6
Domain 3 Task 2

77

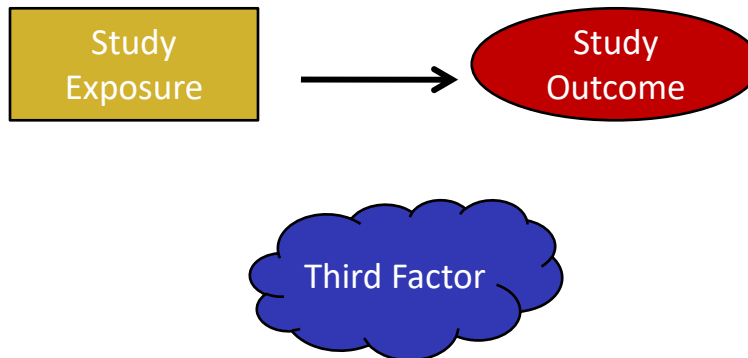
Statistics – Confounders

- Form of bias
- One aspect of the subject has not been separated from another and confounds it, producing a random result
- Describes a true association, but misleading
- In order for a variable to be a confounder, a factor must be associated with exposure and outcome

78

Statistics – Confounders

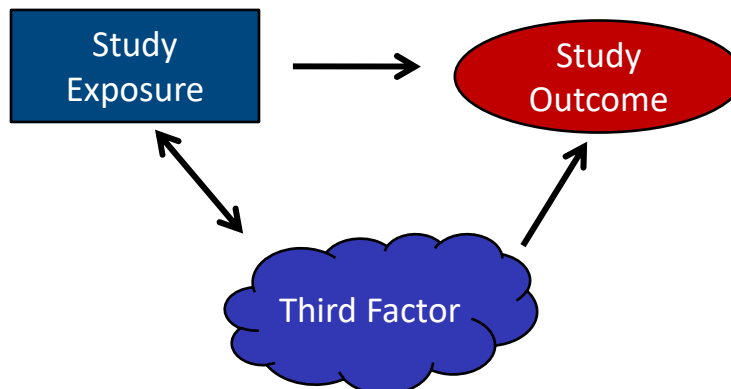
In the perfect study.....



79

Statistics – Confounders

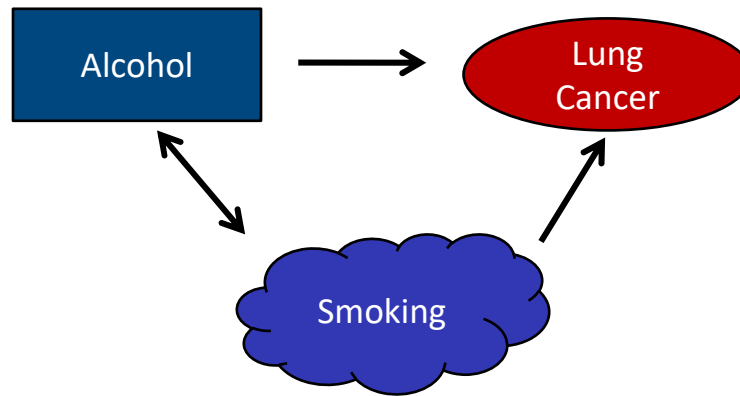
But with confounding...



80

Statistics – Confounders

Example of Confounding



81

Question 4

- A study evaluated the effectiveness of a new antibiotic to treat urinary tract infections (UTIs). There were a significantly higher proportion of subjects in the Drug A group that had chronic use of bladder drainage devices. This can be described as:
 - Interval variable
 - Ratio variable
 - Dependent variable
 - Confounder

Learning Objective 6
 Domain 3 Task 2

82

Question 4 – Answer

- There were a significantly higher proportion of subjects in the Drug A group that had chronic use of bladder drainage devices. This can be described as:
 - A. Interval variable
 - B. Ratio variable
 - C. Dependent variable
 - D. Confounder**

Learning Objective 6
Domain 3 Task 2

83

Question 4 – Answer Explained

- There were a significantly higher proportion of subjects in the Drug A group that had chronic use of bladder drainage devices. This can be described as:
 - A. Interval variable
 - B. Ratio variable
 - C. Dependent variable
 - D. Confounder**

Answers A and B are incorrect: Chronic use of bladder drainage devices is a categorical variable, not an interval or ratio variable.
Answer C is incorrect: In regression analyses, the dependent variable is the outcome variable, in this case microbiological cure.
Answer D is correct: The increased proportion of urinary catheters in the Drug A group is a confounder, a variable that impacts the variables under study.

Learning Objective 6
Domain 3 Task 2

84

Statistics – Confounders

- How to control for confounding
 - Randomize subjects
 - To control for unknown risk factors
 - Inclusion / exclusion criteria
 - Exclude subjects with the confounder
 - Matching subjects
 - Use multivariable analyses that adjust for confounders
 - Adjust for common confounders that may affect your outcome

85

Statistics – Confounders

- Bivariate: One predictor & one outcome (e.g., t-test)
- Multivariable: More than one predictor/more than one outcome at the same time (e.g., regression)
 - Multivariate: 2 or more dependent outcomes or variables; usually for longitudinal data where an outcome is measure for the same individual at multiple time points (repeated measures), modeling of nested/clustered data where there are multiple individuals in each cluster
 - Multivariable: Multiple variables on the right side of the model equation (response or independent variables)

86

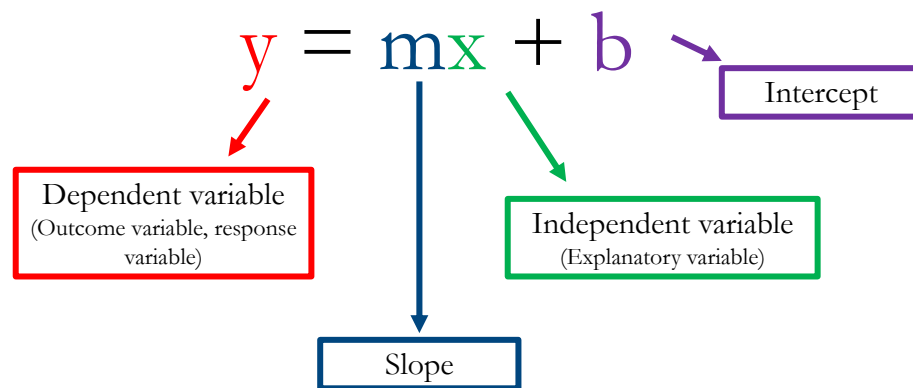
Statistical Analysis

- Bivariate analyses (e.g, t-test):
 - Compare the data
 - Evaluate differences between groups
- Multivariable analyses (e.g, regression):
 - Seek a relationship
 - Does one measurement predict the other?
 - Does one measurement depend on the other?
 - Is there a trend between the two sets of measurements?

87

Statistics – Regression

In algebra.....



88

Statistics – Regression

Dependent variable
(Outcome variable,
response variable)

In algebra.....

$$y = mx + b$$

Slope

In regression.....

Independent
variable
(Explanatory variable)

$$E(y) = \beta_0 + \beta_1 x$$

Intercept

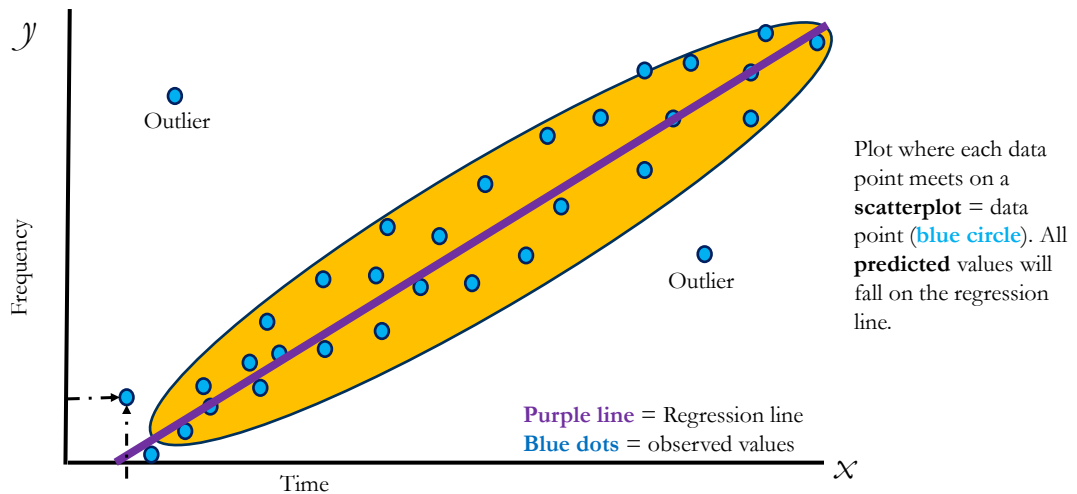
This is the 'model'.

Statistics – Linear Regression

- Dependent variable (outcome) is continuous and independent variables are continuous or categorical
- Assumptions:
 - The correlation, r , is moderate to strong (typically beyond 0.50 or -0.50).
 - The scatterplot must form a linear pattern.
- Steps to linear regression analyses:
 - Determine correlation
 - Plot on scatterplot and verify linear relationship
 - Determine independent variable in the simple linear regression model (usually $p < 0.1$ on bivariate analysis)

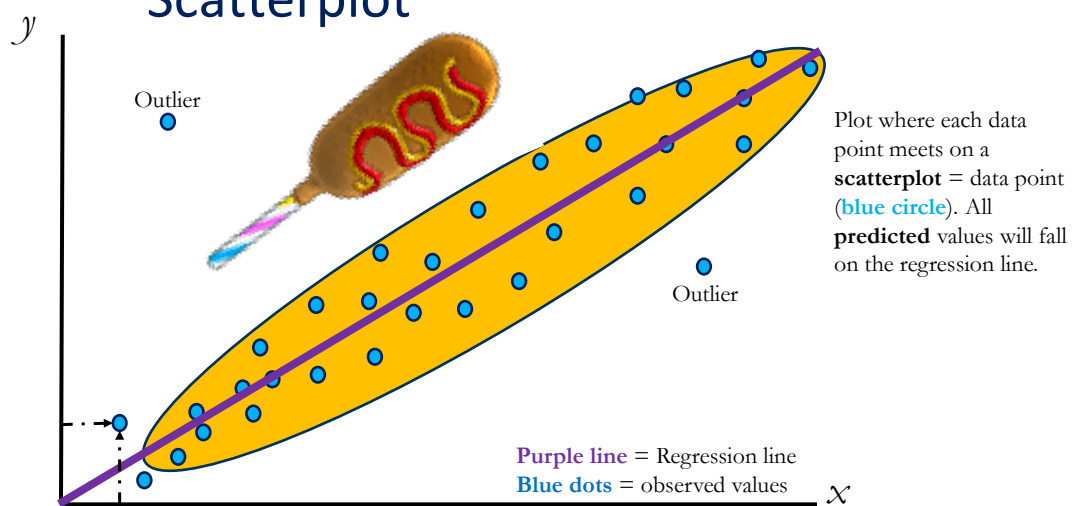
Statistics – Linear Regression Scatterplot

Continuous outcome



91

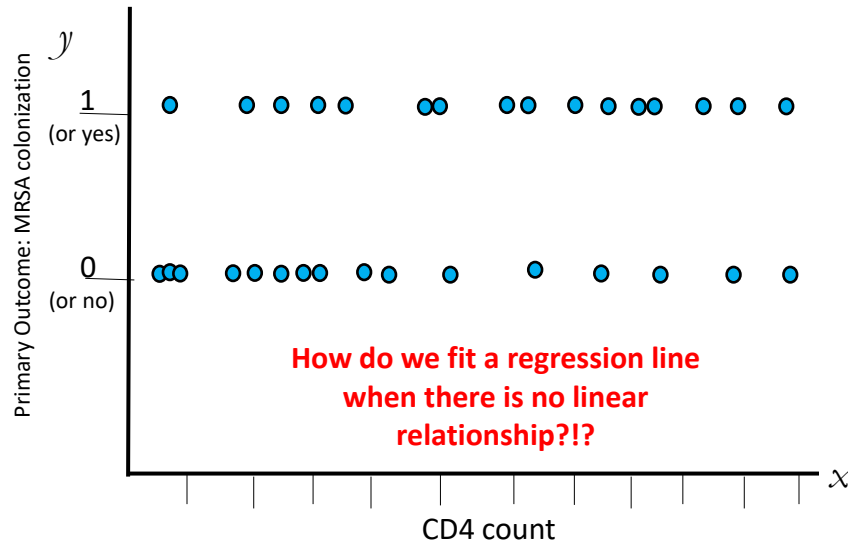
Statistics – Linear Regression Scatterplot



92

Statistics – Logistic Regression

What if your outcome is categorical?



93

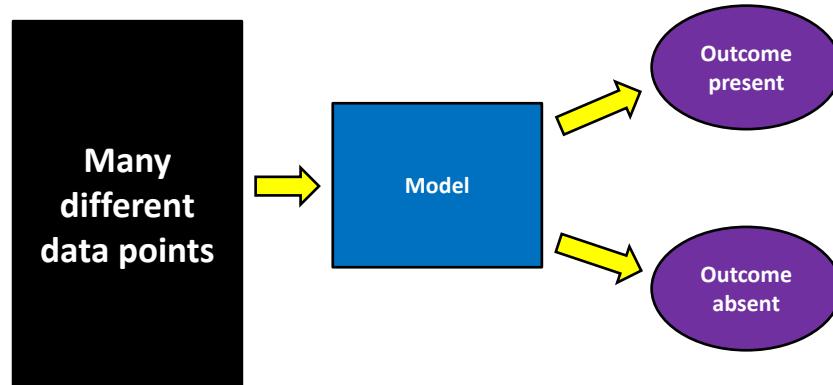
Statistics – Logistic Regression

- Logistic regression:
 - Models the probability of an event occurring depending on the independent variables (categorical or continuous)
 - Estimate the probability that an event occurs
 - Predicts the effect of a series of independent variables on a categorical dependent variable

94

Statistics – Logistic Regression

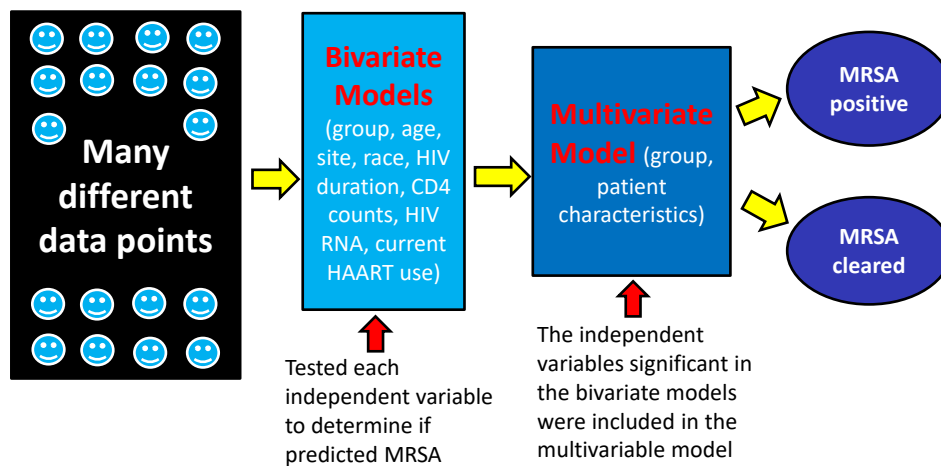
What is the probability that a patient with a value of X would have the outcome of interest?



95

Statistics – Logistic Regression

What are predictors for MRSA clearance at 6 months?



Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

96

Statistics – Simple vs Multiple Regression

- Simple (linear or logistic) regression
 - One independent variable explains the relationship to the dependent variable (a one-to-one relationship)
- Multiple (linear or logistic) regression
 - Multiple independent variables explain the relationship to the dependent variable (a many-to-one relationship)
 - Many variables explain myocardial infarction

97

Statistics – Multiple Regression

- The goal is to select the best independent variables for inclusion in the model and these independent variables are correlated with the dependent variable, but not each other
- The equation in multiple regression is the estimated change in the dependent variable corresponding to a one unit change in a variable when all other variables are held constant
- Cautions with multiple regression analysis:
 - Adequate sample size with multiple independent variables
 - Overfitting: Adding independent variables makes the model worse
 - Multicollinearity: Independent variables are related to each other

98

Statistics – Regression

In algebra.....

$$y = mx + b$$

In simple regression.....

$$E(y) = \beta_0 + \beta_1 x$$

In multiple regression.....

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

Dependent variable
(Outcome variable,
response variable)

Slope

Independent
variable
(Explanatory variable)

Intercept

99

Statistics – Multivariable

Dependent Variable	Independent Variable	# of Independent Variables	Test
Continuous	Continuous	1	Simple Linear regression
Continuous	Continuous	≥ 2	Multiple Linear regression
Continuous	Categorical	1	Simple Linear regression
Continuous	Categorical	≥ 2	Multiple Linear regression
Categorical	Continuous	1	Simple Logistic regression
Categorical	Continuous	≥ 2	Multiple Logistic regression
Categorical	Categorical	1	Simple Logistic regression
Categorical	Categorical	≥ 2	Multiple Logistic regression

100

Statistics – Multivariable

Dependent Variable	Independent Variable	# of Independent Variables	Test
Continuous	Continuous	1	Simple Linear regression
		≥ 2	Multiple Linear regression
Continuous	Categorical	1	Simple Linear regression
		≥ 2	Multiple Linear regression
Categorical	Continuous	1	Simple Logistic regression
		≥ 2	Multiple Logistic regression
Categorical	Categorical	1	Simple Logistic regression
		≥ 2	Multiple Logistic regression

'Simple' Model

Multivariate Model

101

Question 5

- Which statistical method could best adjust the primary study outcome (microbiological cure) to control for the difference in urinary catheter use in the previously mentioned study?
- Logistic regression
 - Chi Square
 - ANCOVA
 - Mann-Whitney U test

Learning Objective 2
 Domain 3 Task 2

102

Question 5 – Answer

- Which statistical method could best adjust the primary study outcome (microbiological cure) to control for the difference in urinary catheter use in the previously mentioned study?

- A. Logistic regression**
- B. Chi Square
- C. ANCOVA
- D. Mann-Whitney U test

Learning Objective 2
Domain 3 Task 2

103

Question 5 – Answer Explained

- Which statistical method could best adjust the primary study outcome (microbiological cure) to control for the difference in urinary catheter use in the previously mentioned study?

- A. Logistic regression**
- B. Chi Square
- C. ANCOVA
- D. Mann-Whitney U test

Answer A is correct: This study outcome being evaluated is categorical; dependent variables in logistical regression are categorical and allow for continuous or categorical independent covariates.

Answers B and D are incorrect: Neither Chi square nor the Mann-Whitney U test allow for adjustment with a covariate.

Answer C is incorrect: While ANCOVA does allow for categorical independent covariates, the dependent variable is continuous.

Learning Objective 2
Domain 3 Task 2

104

Statistics – Time-To-Event Analysis

- Also called survival analysis
- Follows a group of participants until an “event” occurs
 - Event: death, hospitalization, development of disease
- Takes into account dropouts, different study entry time and different follow-up time
- Within a study period, these analyses determine:
 - The frequency of participants that experience the event
 - The time period at which the event occurs
 - Variables (risk factors) that influence the likelihood of an event
 - How treatment affects time to an event

105

Statistics – Time-To-Event Analysis

- There are typically 3 components of time-to-event analysis:
- **Kaplan Meier Curve**
 - Estimates event probabilities for each group which is shown in a **descriptive graph**
 - Participants are censored when the event occurred OR at the end of follow-up
- Logrank test
 - Determines if the Kaplan Meier curves are **different**
- Cox proportional hazard regression analysis
 - Determines differences in the **time to event** occurring for each group adjusting for various **covariates**

106

Statistics – Time-To-Event Analysis

Kaplan Meier was done for:

- 1) Time to MRSA clearance using monthly swab data
- 2) Effectiveness of study drug on subsequent MRSA infection/SSTI

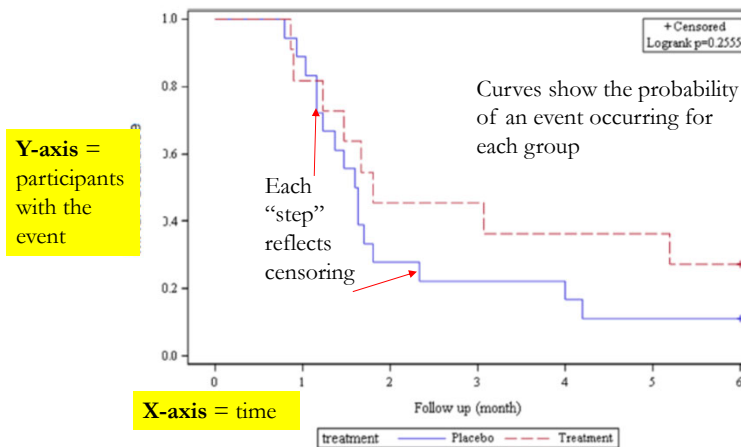


Fig 2. Kaplan-Meier Curve for MRSA Clearance among Randomized Participants.

Image adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

107

Statistics – Time-To-Event Analysis

- There are typically 3 components of time-to-event analysis:
- Kaplan Meier Curve
 - Estimates event probabilities for each group which is shown **graphically**
 - Participants are censored when the event occurred OR end of follow-up
- Logrank test
 - Determines if the Kaplan Meier curves are **different**
- Cox proportional hazard regression analysis
 - Determines differences in the **time to event** occurring for each group adjusting for various **covariates**
 - Reports a **Hazard Ratio**

108

Statistics – Time-To-Event Analysis

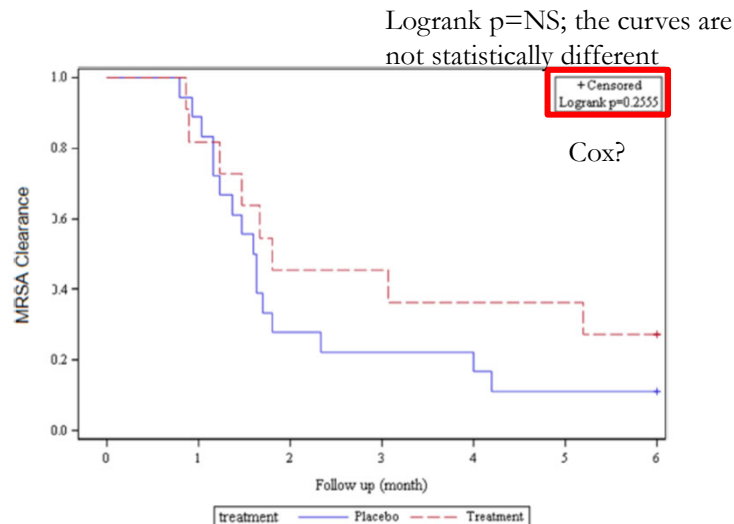


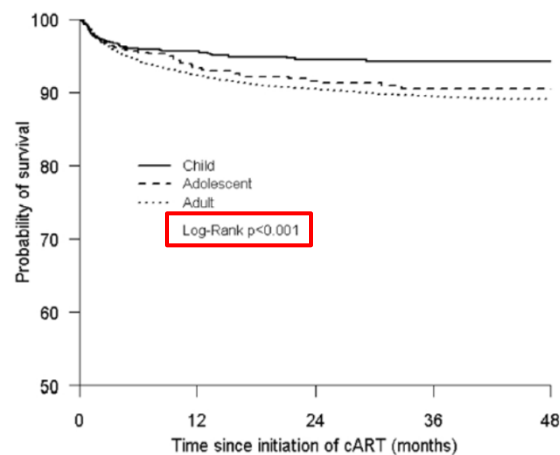
Fig 2. Kaplan-Meier Curve for MRSA Clearance among Randomized Participants.

Image adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

109

Statistics – Time-To-Event Example

Different example:
time to mortality
after ART initiation



# at risk for death	Number at Risk					Events		# of deaths
	0	12	24	36	48	0	48	
—	810	757	488	249	60	44		
- - -	575	517	320	168	45	49		
.....	21982	19831	14435	8137	4806	2201		

Image adapted under the Creative Commons license. Bakanda C, et al. Survival of HIV-infected adolescents on antiretroviral therapy in Uganda: Findings from a nationally representative cohort in Uganda. *PLoS One*. 2011;6:e19261.

110

Was an appropriate statistical method used?



111

Statistics used in Weintrob, et al



- Fisher's exact test - Primary outcome
- GEE with a logit link function - Primary outcome which included all participants time points – accounted for subject level correlation
 - Is a type of linear regression analysis
- Logrank test - Secondary analysis – time to MRSA clearance using monthly swab data and effectiveness of study drug on subsequent MRSA infection/SSTI
- Time-to-event Cox proportional hazards models - Secondary analysis – time to MRSA clearance using monthly swab data and effectiveness of study drug on subsequent MRSA infection/SSTI
- Simple and multivariable logistic regression
 - To evaluate for predictors (including demographics, clinical factors) for MRSA clearance at the 6-month time point

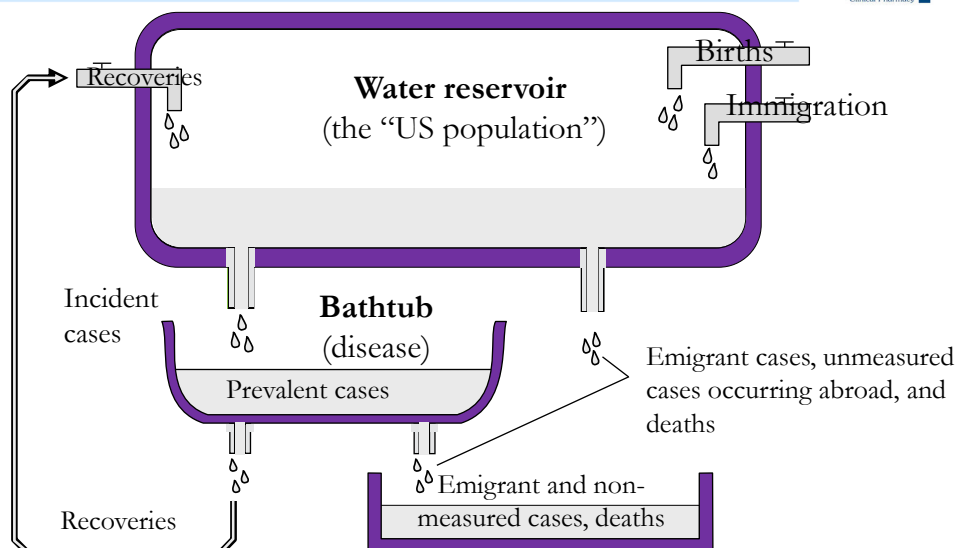
Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

112

Epidemiologic Measurements

113

Incidence vs. Prevalence



114

Contingency Table (2x2 Table) Results

		Event?	
		Yes	No
Exposure?	Yes	A	B
	No	C	D

115

Example (fake #s):

- Oseltamivir vs placebo in influenza-like illness-related hospitalization

		Event		
		Hospitalization	No hospitalization	Total
Exposure	Oseltamivir	50 (5%)	950	1000
	Placebo	75 (7.5%)	925	1000
	Total	125	1875	2000

116

Absolute Risk

- Actual number of events prevented
- “X number of extra cases in 100, 1000, or 10,000 people”
 - 75/1000 hospitalizations in the placebo group vs 50/1000 in the oseltamivir group
 - $75 - 50 = 25$ extra cases of hospitalization per 1000 people

117

Relative Risk (RR)

- Compares the chance of an event between groups
- Is the best estimate of the strength or magnitude of the association between exposure and outcome

118

Relative Risk

		Event?	
		Yes	No
Exposure?	Yes	A	B
	No	C	D

$$RR = \frac{(A/A+B)}{(C/C+D)}$$

This can also be translated as the risk of the outcome in the exposed (the 'Yes' row in the table) OVER the risk of the outcome in the non-exposed (the 'No' row)

119

	Hospitalization	No hospitalization	Total
Oseltamivir	50 (5%)	950	1000
Placebo	75 (7.5%)	925	1000
Total	125	1875	2000

$$RR = \frac{(A/A+B)}{(C/C+D)} = \frac{(50/50+950)}{(75/75+925)} = \frac{0.05}{0.075} = 0.67$$

$$RR = 0.67$$

Relative Risk Reduction (RRR): $(1-RR)*100\% = (1-0.67)*100$
 33% decreased risk of hospitalization in the oseltamivir group

120

Odds Ratio

- Used in case control studies to estimate risk
- Cannot evaluate the impact of the risk....the event is more or less likely via odds
- Overestimates RR

121

Odds Ratio

		Event?	
		Yes	No
Exposure?	Yes	A	B
	No	C	D

$$OR = \frac{(A/C)}{(B/D)}$$

122

	Hospitalization	No hospitalization	Total
Oseltamivir	50 (5%)	950	1000
Placebo	75 (7.5%)	925	1000
Total	125	1875	2000

$$OR = \frac{(A/C)}{(B/D)} = \frac{(50/75)}{(950/925)} = 0.65$$

OR = 0.65; 35% decreased **odds** of hospitalization in the oseltamivir group

123

Interpretation of RR and OR

Table. Direction of Risk Associated with OR and RR

RR	OR	Interpretation
< 1	< 1	Negative association (protective) RR: Risk of disease is lower in the exposed group OR: Odds of exposure is lower in the diseased group
= 1	= 1	No association (disease is equally likely) RR: Risk of disease in the two groups is the same OR: Odds of exposure in the two groups is the same
> 1	> 1	Positive association RR: Risk of disease is greater in the exposed group OR: Odds of exposure is greater in the diseased group

Table. Magnitude of Risk Associated with OR and RR

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

124

Question 6

- A case-control study is conducted to determine whether proton pump inhibitor (PPI) use is associated with an increased risk of developing *C. difficile* infection (CDI). The final analysis shows the odds ratio (OR) for CDI with PPI exposure to be 1.3 (95% confidence interval [CI], 0.8–1.5). Which best describes the results?
 - A. PPI exposure increases the risk of CDI by 130%.
 - B. PPI exposure reduces the risk of CDI by 20%.
 - C. PPI exposure increases the risk of CDI by 30%.
 - D. PPI exposure is not associated with an increased risk of CDI.

Learning Objective 7
Domain 3 Task 2

125

Question 6 – Answer

- A case-control study is conducted to determine whether proton pump inhibitor (PPI) use is associated with an increased risk of developing *Clostridium difficile* infection (CDI). The final analysis shows the odds ratio (OR) for CDI with PPI exposure to be 1.3 (95% confidence interval [CI], 0.8–1.5). Which best describes the results?
 - A. PPI exposure increases the risk of CDI by 130%.
 - B. PPI exposure reduces the risk of CDI by 20%.
 - C. PPI exposure increases the risk of CDI by 30%.
 - D. PPI exposure is not associated with an increased risk of CDI.**

Learning Objective 7
Domain 3 Task 2

126

Question 6 – Answer Explained

- A case-control study is conducted to determine whether proton pump inhibitor (PPI) use is associated with an increased risk of developing Clostridium difficile infection (CDI). The final analysis shows the odds ratio (OR) is 1.3 (95% confidence interval [CI], 0.8–1.5). Which of the following is the correct interpretation of the results?
- A. PPI exposure increases the risk of CDI by 30%.
 - B. PPI exposure reduces the risk of CDI by 20%.
 - C. PPI exposure increases the risk of CDI by 1.3 times.
 - D. PPI exposure is not associated with an increased risk of CDI.**

Answer D is correct; Answers A–C are incorrect. A correct interpretation of the results is recognizing that even though the OR suggests an associated increase of 30% in the risk of being exposed to CDI, the 95% CI crosses 1, meaning that the odds of exposure to a PPI are as likely to increase the risk as to decrease the risk of developing CDI.

Learning Objective 7
 Domain 3 Task 2

127

Absolute Risk Reduction (ARR)

- Subtract the incidence in the outcome of interest between groups
- Provides an overall estimate of the decrease in risk
 - 7.5% hospitalization in placebo group vs. 5% hospitalization in oseltamivir group
 - $7.5 - 5 = 2.5\%$ ARR of hospitalization in patients taking oseltamivir
- Better for clinical decision making vs RRR
 - RRR does not take baseline risk into account

128

Number Needed To Treat (NNT)

- NNT helps to put the results in perspective
 - Are the benefits of treatment worthwhile?
- Number of patients that need to be treated to benefit one patient
- Less likely to be misleading because calculated on absolute risk
- $1 / \text{ARR}$
- Oseltamivir and hospitalization:
 - $1 / 2.5\% = 1 / .025 = 40$ (always round up)
 - 40 patients needed to be treated in order to prevent 1 hospitalization

129

Importance Of NNT And NNH

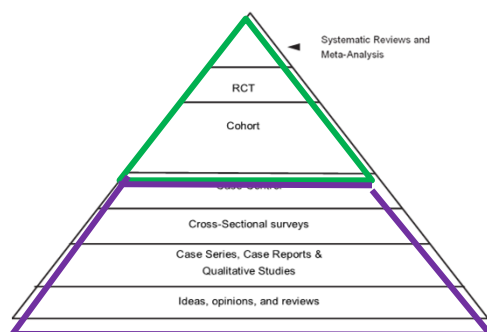
- Comparing these two numbers can help you evaluate the risks and benefits of treatment
- Example: Antibiotics for the treatment of otitis media-related pain in pediatric patients
 - $\text{NNT} = 15 = 15$ patients need to be treated with antibiotics to relieve pain in one child
 - $\text{NNH} = 12 = 12$ patients need to be treated with antibiotics to have a side effect
- You can calculate!

130

Additional Study Designs

131

Study design - Hierarchy of evidence



Determine
cause and
effect
relationships
(causal)

Determine
associations

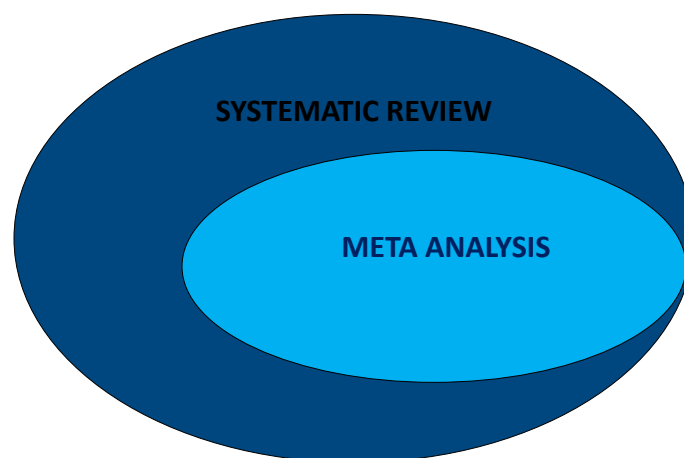
132

Systematic Reviews

- Method used to summarize and combine medical literature in a structured manner
- Makes clinical information more useful to clinicians
- Quality systematic reviews should:
 - Use explicit methods
 - Have a focused question(s)
 - Be objective
 - Be reproducible

133

Systematic Review vs Meta - Analysis



134

Meta-analysis

- Combines results from several different studies to form conclusions
 - Usually use results from RCT
- Increases the certainty of results
- Useful in rare diseases, variation in results between studies, and when several small studies are available that had a small sample size
- Issues:
 - Quality of original study design
 - Quality of literature search
 - Differences of inclusion / exclusion criteria
 - Differences in treatment dosages, control used, and patient demographics between studies

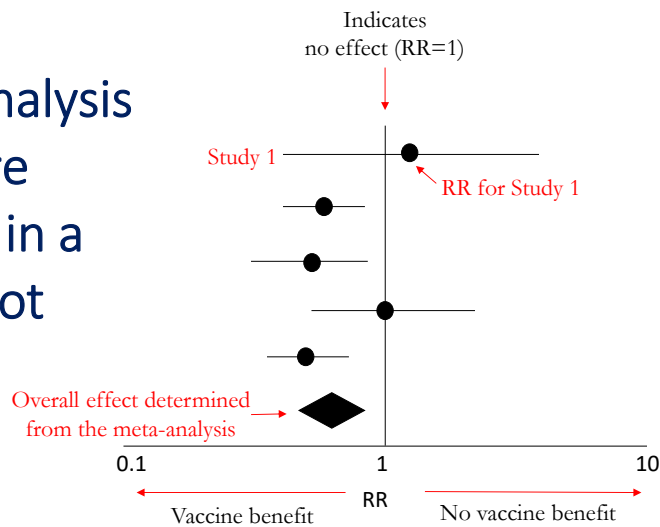
135

Meta - Analysis

- Easier to understand than all the individual studies
- Advantages:
 - Small n
 - Rare diseases
 - Results differ
- Disadvantages
 - Comparing apples to apples?
 - Search for studies
 - Publication bias
 - Quality of the studies?

136

Meta – Analysis Results are Reported in a Forrest Plot



137

Systematic Review vs Meta Analysis

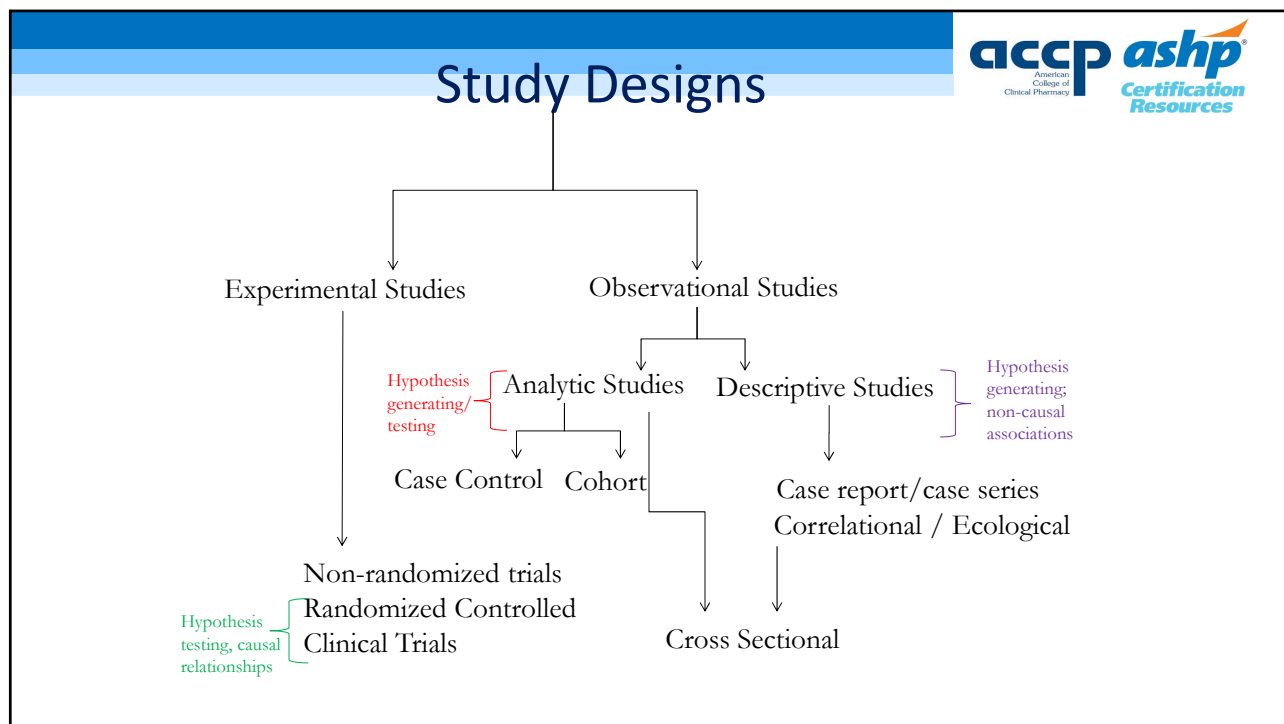
Systematic Review

- Basis is multiple clinical trials
- Studies derived from an extensive and structured search of the literature
- Explicit study inclusion/exclusion criteria
- Validity of studies assessed


Meta Analysis

- All elements of systematic review
- **Also:**
 - All studies are statistically combined
 - Individual study results are integrated into a single result
- Not always possible

138



139



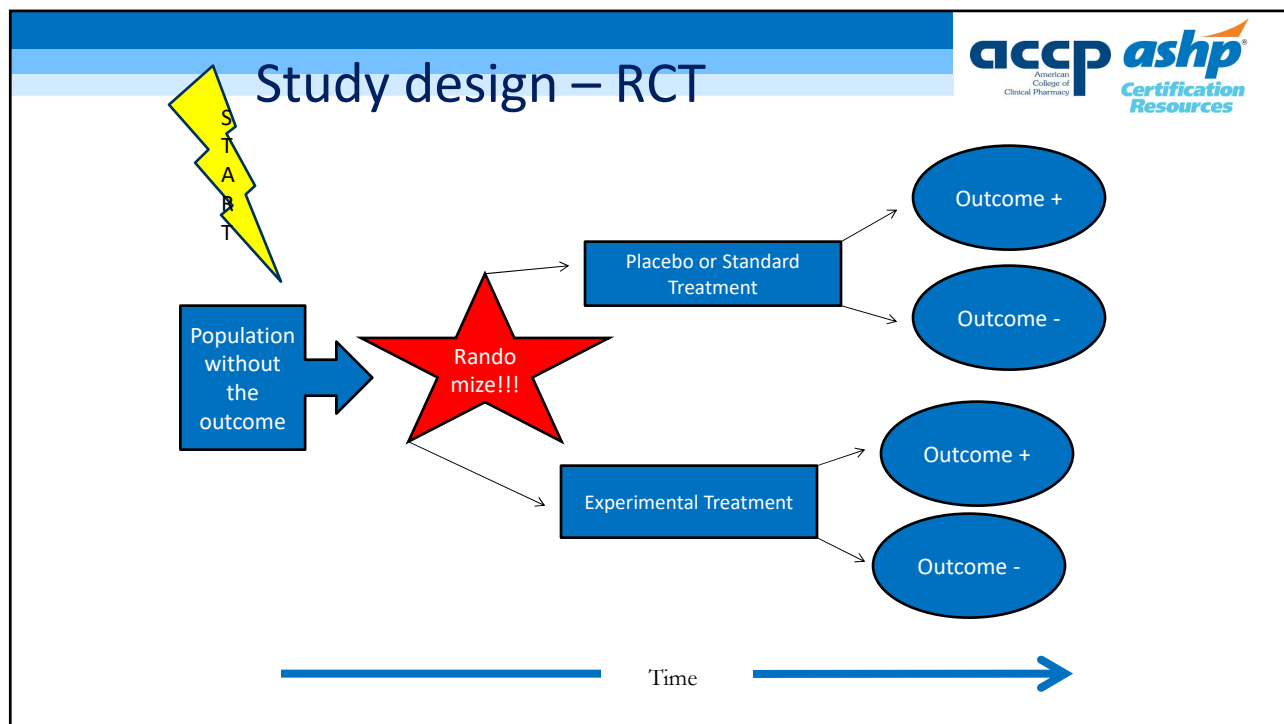
 American College of Clinical Pharmacy

 Certification Resources



Study design – RCT

- Randomized controlled clinical trial (RCT)
- Prospective, experimental, hypothesis testing
- Patients are randomly assigned to:
 - Intervention – “experimental” treatment
 - Control – “nonexperimental” treatment
- Most reliable results that are likely to impact patient care
- Want the subjects to be blinded:
 - Single-blind: subjects do not know treatment assignment
 - Double-blind: neither the subjects nor the investigators know treatment assignment
 - Triple-blind: neither the subjects, investigators, or the data analyst know the treatment assignment
- **GOLD STANDARD** for studying interventions

140



141

Study design – Parallel RCT

- Parallel designed RCT
- Majority of RCT
- Separate patients into two or more groups
 - One group control, the other receives the active intervention

Nasal mupirocin + soap (intervention)

Placebo (control)

END (and/or randomized again if positive)

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

142

RCTs: Superiority vs. Equivalence vs. Non-inferiority

- **Superiority:** Detect a difference between treatments
 - Typical design in a clinical trial.
- **Equivalence:** Confirm the absence of meaningful difference(s) between treatments
 - What difference is important?
- **Non-inferiority:** Investigate whether a treatment is not clinically worse (no less effective)
 - May be the most effective, or have a similar effect.
 - If placebo is not possible due to ethical reasons

143

RCTs: Superiority vs. Equivalence vs. Non-inferiority

- **Non-directional hypothesis:**
 - **Difference:** Are the means different?
 - $H_0: \text{Mean}_1 = \text{Mean}_2$
 - **Equivalence:** Are the means practically equivalent?
 - $H_0: \text{Mean}_1 - \text{Mean}_2 \geq \Delta$
- **Directional hypothesis:**
 - **Superiority:** Is mean 1 > mean 2?
 - $H_0: \text{Mean}_1 \leq \text{Mean}_2$
 - **Non-inferiority:** Is mean 1 no more than a certain amount lower than mean 2?
 - $H_0: \text{Mean}_1 - \text{Mean}_2 \geq \Delta$

144

Cohort Study

- Observational, hypothesis generating or hypothesis testing
- Utilizes a defined study population
 - Start with a “Healthy Cohort”
- Study groups are defined by their exposure and then the outcome is compared (usually)
 - Determine exposure (or potential risk factors) before the outcome occurs
 - Determine outcome identically in both groups
- Subjects are followed over time to attempt to correlate risk factors and disease occurrence

145

Prospective Cohort Study

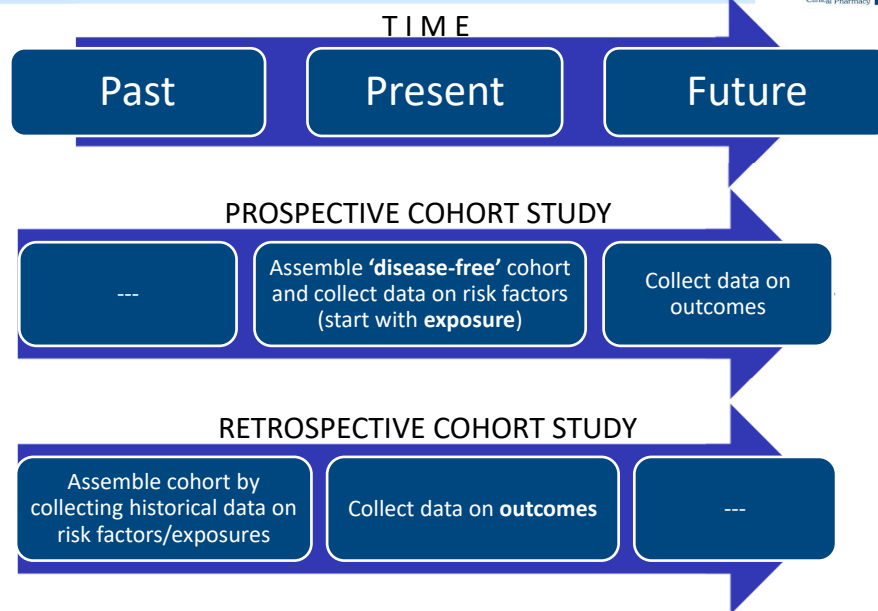
- Majority of cohort studies
- Study groups start in the present time and baseline data is collected and continues for a period of time
- Outcomes have not occurred at baseline
- Starts with evaluating exposure, then the outcome of interest
- ‘Truer’ estimate of outcomes and risk factors

146

Retrospective Cohort Study

- Historical group is defined and present outcomes (or events) are evaluated
 - Thus, the outcomes have already occurred at baseline
- May miss disease of short duration or fatal cases
- Decreases time and cost of a prospective study
- Unable to monitor and control data collection / standardize care
- Not preferred
 - Less complete and accurate
 - But can be used if good, comprehensive records are available

147



148

Case Control Study

- Observational, hypothesis generating or “hypothesis testing”
- Study groups are selected on the basis of the outcome and evaluates past exposure to possible risk factors
 - “Having the disease of interest versus not having the disease of interest”
- Control group is matched to the case group individually based on common characteristic(s) from the same setting (geographic area, hospital, health care provider)
 - Two or more control groups may be used to attempt to decrease bias
- Data on potential exposure to risk factors is collected by medical record review or by interviewing the subject
- Actual risk of the outcome cannot be determined
 - Estimate of the risk = odds ratio

149

Case Control Study

- The best place for a case control study:
 - Need a study completed quickly and/or inexpensively
 - Rare disease (<1%) or disease of long latency
 - But need documentation of disease occurrence
 - Need evidence to generate a hypothesis especially when investigating the causes of a disease

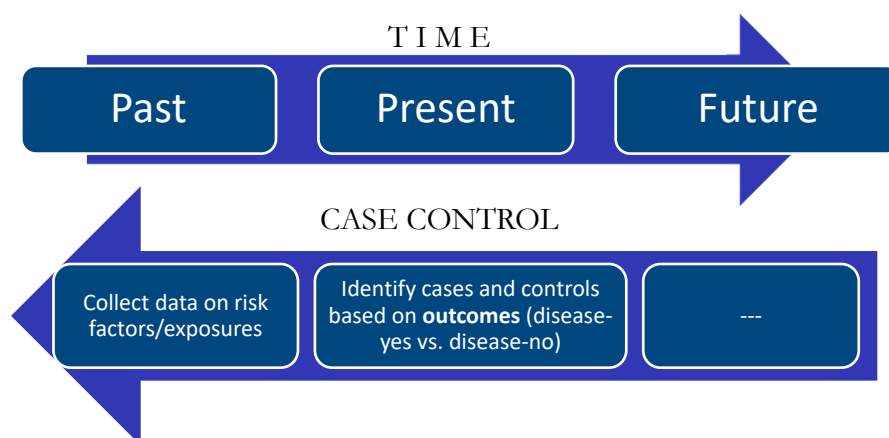
150

Case Control Study

- Select cases (w/ disease) from a source population
 - Study cases should be representative of all cases in the population
 - Inclusion criteria should be independent of exposure status
 - Clear diagnostic criteria to define cases
- Select controls (w/o disease)
 - From the same population at risk as the cases
 - Participation does not depend on exposure
- Need equivalent reliability of the information in both groups
- Exposure status (or risk factors) are unknown at baseline
- Determine exposure status retrospectively in both groups
- Analyze results for differences in exposure between groups

151

Case Control Study Design



152

Matching

- Match controls to cases to control for confounding
- Use patient demographics
 - Age, sex, location, time period of disease
 - Use factors that are known to greatly influence the outcomes but do not pertain to the question being asked
- Can use more than one control for each case
 - Increases power
 - Especially to study very rare outcomes
- Has pros and cons.....but most investigators do it if possible.

153

Case Control Study – Selecting Controls

CASES	CONTROLS
All cases diagnosed in the community	Sample of the general population
All cases diagnosed in all hospitals	Sample of patients in all hospitals who do not have the disease
All cases diagnosed in a single hospital	Sample of patients in the same hospital who do not have the disease

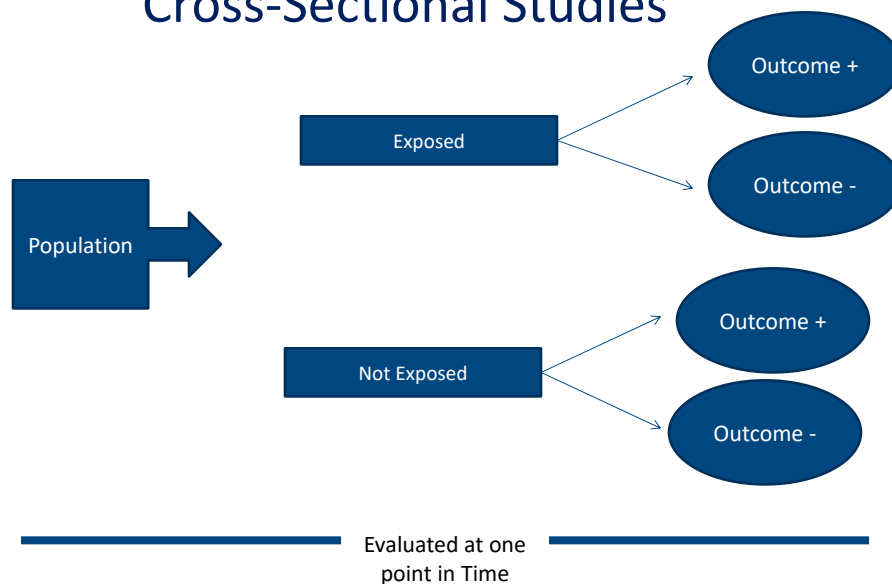
154

Cross-Sectional Studies

- Observational, hypothesis-generating
- Snapshot of a population at one point in time
- Most surveys and mass screening programs are cross-sectional surveys
- Appropriate for:
 - Determining the prevalence of risk factors
 - Frequency of prevalent cases of diseases for a defined population
 - Evaluating current health status and future planning
- Classic example: US census
- Can repeat to determine changes in risk factors and changes in disease frequency over time

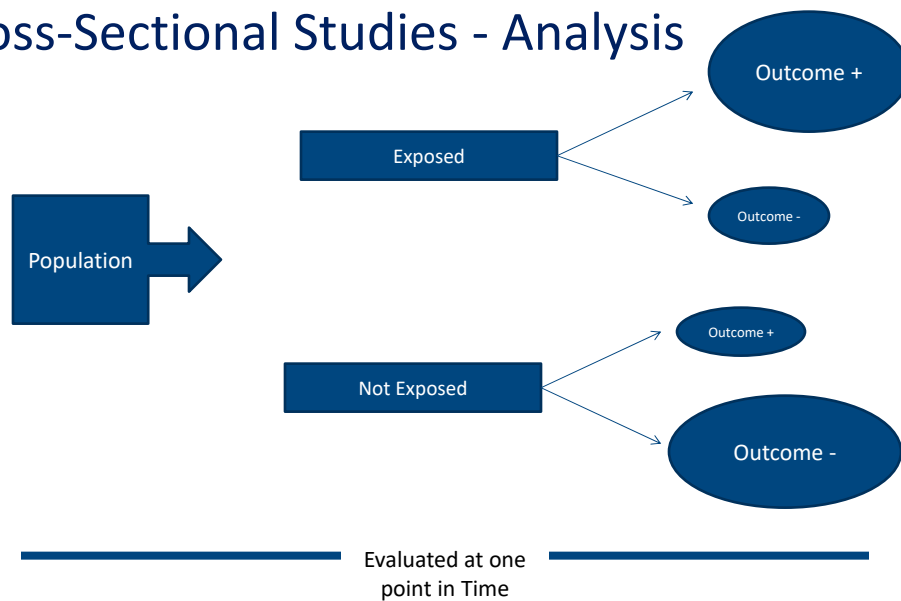
155

Cross-Sectional Studies



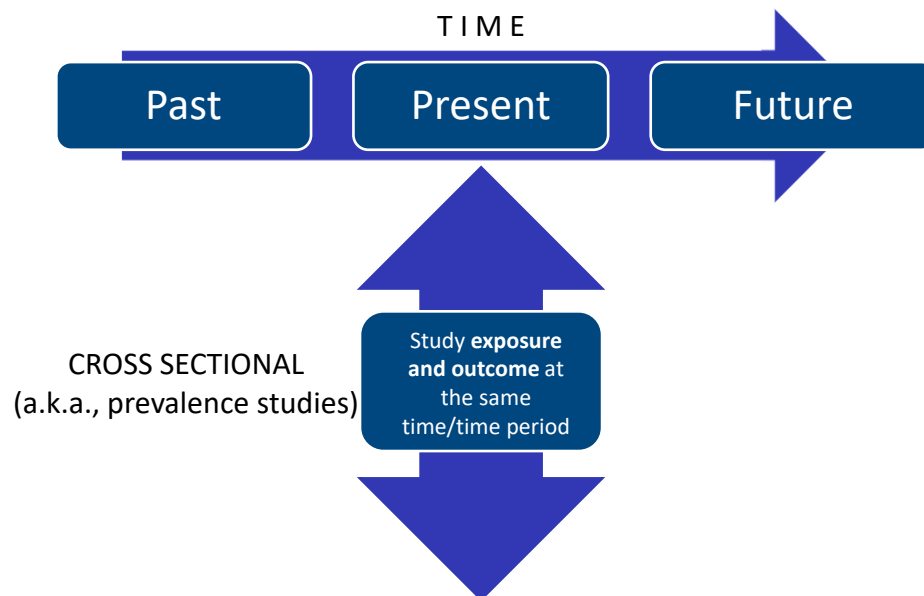
156

Cross-Sectional Studies - Analysis



157

Cross-Sectional Study Design



158

Observational Study Designs – Summary of Characteristics

	Case Report/ Series	Case-control	Cohort	Cross-sectional
Measure of Association		OR	RR, HR	Prevalence
Description		Persons exposed to risk are compared to those not exposed	Persons with Dz compared to those without Dz	Exposure and Dz are examined in a sample at one point
Major Advantages	<ul style="list-style-type: none"> Generate new information about natural history of Dz ID new disease/ condition 	<ul style="list-style-type: none"> Study relatively rare diseases Low cost and short duration 	<ul style="list-style-type: none"> Study relatively rare exposures Study temporal associations Direct risk estimates 	<ul style="list-style-type: none"> Low cost and short duration
Major Disadvantages	<ul style="list-style-type: none"> Usually can't measure rates of association 	<ul style="list-style-type: none"> Not practical for studying rare <u>exposures</u> Inability to study multiple outcomes 	<ul style="list-style-type: none"> Not practical for rare Dzs Increased cost and longer duration (prospective) 	<ul style="list-style-type: none"> Temporal associations can't be established

Adapted from *Pharmacotherapy* 2010;30:973-984

159

Question 7

- A study seeks to determine the impact of adverse drug events on patient outcomes in FDA-approved drugs. Which would be the best approach to conducting this study?
 - A randomized controlled clinical trial (RCT) with a test for continuous variables to determine the difference in outcomes.
 - A retrospective case-control study with a test for proportions to determine the difference in outcomes.
 - A prospective observational study with survival analysis to determine the difference between cohorts.
 - A retrospective cohort study with a test for proportions to determine the difference in outcomes.

Learning Objective 3
Domain 3 Task 2

160

Question 7 – Answer

- A study seeks to determine the impact of adverse drug events on patient outcomes in FDA-approved drugs. Which would be the best approach to conducting this study?
 - A. A randomized controlled clinical trial (RCT) with a test for continuous variables to determine the difference in outcomes.
 - B. A retrospective case-control study with a test for proportions to determine the difference in outcomes.
 - C. A prospective observational study with survival analysis to determine the difference between cohorts.**
 - D. A retrospective cohort study with a test for proportions to determine the difference in outcomes.

Learning Objective 3
Domain 3 Task 2

161

Question 7 – Answer Explained

- A study seeks to determine the impact of adverse drug events on patient outcomes in FDA-approved drugs. Which would be the best approach to conducting this study?
 - A. A randomized controlled clinical trial (RCT) with a test for continuous variables to determine the difference in outcomes.
 - B. A retrospective case-control study with a test for proportions to determine the difference in outcomes.
 - C. A prospective observational study with survival analysis to determine the difference between cohorts.**
 - D. A retrospective cohort study with a test for proportions to determine the difference in outcomes.

Answer A is Incorrect: To effectively determine the incidence and clinical impact of adverse drug events on clinical outcomes, it would be unethical to randomize patients to experience the event.
Answers B and D are Incorrect: A retrospective design would not be ideal because of the limitations in data extraction, assignment of events, and interpretation of causality.
Answer C is Correct. A prospective observational design would allow the investigator team to identify the incidence of events and sequential events and determine causality.

Learning Objective 3
Domain 3 Task 2

162

Key Takeaways

163

Trial Design and Biostatistics

2022 ACCP/ASHP Ambulatory Care Preparatory Review and Recertification Course

Katie J. Suda, PharmD, M.S.

Research Health Scientist and Associate Director of Clinical Therapeutics
VA Center for Health Equity Research and Promotion

Professor of Medicine and Pharmacy and Therapeutics
Director, Transition to Independence Program
University of Pittsburgh Schools of Medicine and Pharmacy
Pittsburgh, Pennsylvania

164

30 Evaluative Questions To Ask of a Trial

Acknowledgement:

Dr. Lee Vemeulen

University of Kentucky

UK Healthcare



165

30 Evaluative Questions

- Systematic approach to clinical trial evaluation
- Not all questions necessarily of equal weight
- Many questions overlap
- Use as tool to find important flaws in clinical research
- No perfect trial
- Goal: become an informed consumer and use high quality evidence
- Included in your handout



166

Overview

- 1. What is the journal's reputation? Are the articles refereed?
- 2. Is the title or abstract misleading? Does the author's bias show?
- 3. Are the researchers qualified to undertake this study?
- 4. Is the location of the study adequate or appropriate?
- 5. Is the article referenced with key up-to-date articles?

167

Introduction & Study Design

- Introduction
- 6. Is there a brief review of previous work and background on why the study was done?
- 7. Is the hypothesis or objectives of the trial clearly stated?
- Overall Study Design
- 8. Was the study design appropriate to the hypothesis?
Ethical?

168

Methodology

- 9. Was subject selection and exclusion criteria clearly detailed? Was subject selection adequate for extrapolation to the appropriate population?
- 10. Was the number of subjects enrolled adequate?
- 11. Was the subject sample described sufficiently (age range, sex, severity of disease) to make sure groups were comparable? Were group differences subjected to statistical analysis?
- 12. Were appropriate controls used?
- 13. Was allocation to treatment groups truly random?
- 14. Were doses, schedules and duration of drug treatment adequate and comparable?

169

Methodology

- 15. Were washouts used? Were they of sufficient duration?
- 16. Was concurrent therapy allowed? Controlled?
- 17. Was the study blinded? Was it truly blinded?
- 18. Were observers identified? Were they qualified? Were they blinded?
- 19. Were the test measures used indicative of therapeutic efficacy? Could a more reliable test measurement have been used?
- 20. Did the test measurements use subjective or objective assessment?
- 21. How long were subjects followed? Was it long enough?

170

Results

- 22. Were the results clearly, accurately and adequately presented?
- 23. Were all the results presented?
- 24. Were dropouts adequately accounted for?
- 25. Was the impact of patient compliance considered?
- 26. Was an appropriate statistical method used?
- 27. Does statistically significant difference necessarily imply clinical significance?

171

Discussion and Conclusion

- 28. Were valid conclusions based upon the results presented?
- 29. Were valid conclusions based upon the hypothesis of the study?
- 30. Does the discussion place the results of this study into the perspective of previous clinical trial comparing and contrasting results? Does the discussion honestly outline the clinical trials shortcomings?

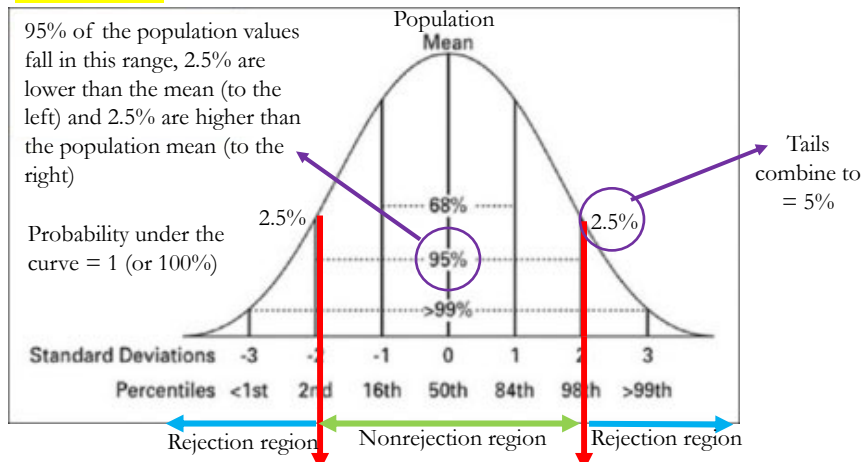
172

Additional Decision Error Slides

173

Type I Error Visualized

Alpha = 0.05



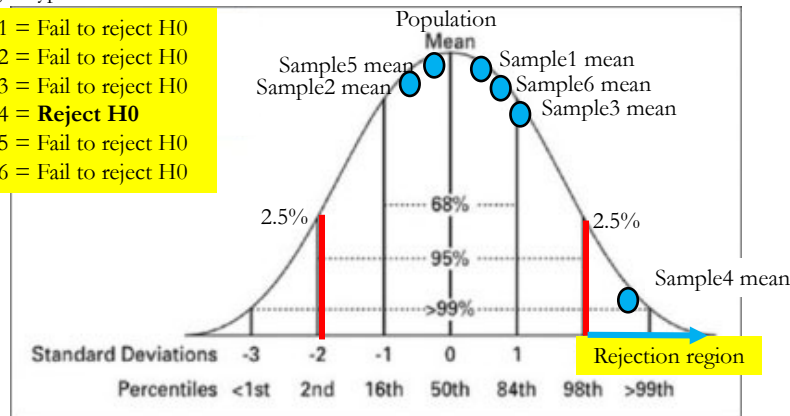
The Critical Values for a two-tailed test. A two-tailed test indicates that the investigator is unsure where the study result will fall. Outside these areas (towards the tails) is the rejection area.

174

Type I Error Visualized

Alpha = 0.05, the level of tolerance for making a Type I Error

Sample1 = Fail to reject H0
Sample2 = Fail to reject H0
Sample3 = Fail to reject H0
Sample4 = **Reject H0**
Sample5 = Fail to reject H0
Sample6 = Fail to reject H0



If we took a sample and it was by chance like Sample4 we would **incorrectly** reject the null hypothesis. This is a **Type 1 Error**.

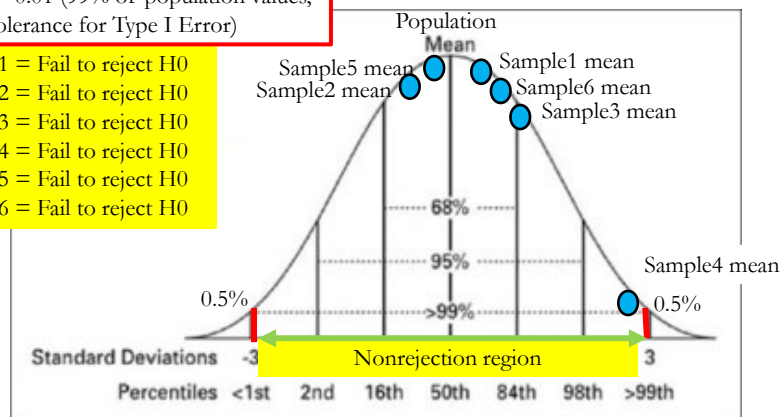
175

Type I Error Visualized

What if.....

Alpha = 0.01 (99% of population values, lower tolerance for Type I Error)

Sample1 = Fail to reject H0
Sample2 = Fail to reject H0
Sample3 = Fail to reject H0
Sample4 = Fail to reject H0
Sample5 = Fail to reject H0
Sample6 = Fail to reject H0



H0 = No difference; the study mean is similar to the population mean, lying between the red bars
H1 = There is a difference

176

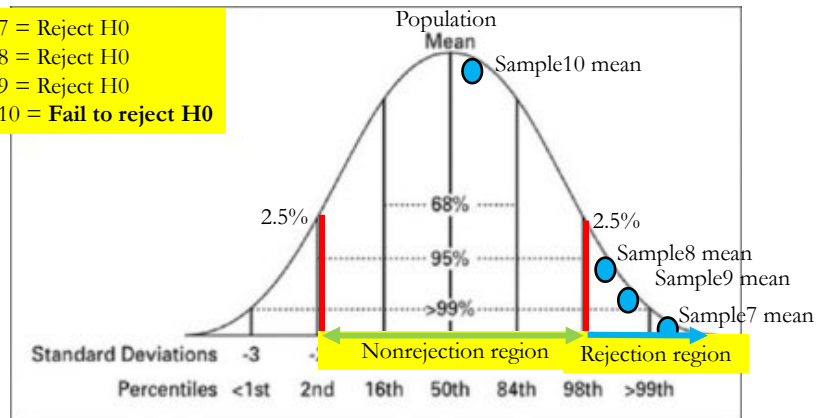
Type II Error Visualized

Alpha = 0.05

Sample7 = Reject H0
Sample8 = Reject H0
Sample9 = Reject H0
Sample10 = **Fail to reject H0**

H0 = No difference; the study mean is similar to the population mean, lying between the critical values

H1 = There is a difference



If we took a sample and it was by chance like Sample10 we would **incorrectly fail to reject (or accept) the null hypothesis**. This is a **Type 2 Error**.

177

Additional Biostatistics Slides

178

Statistics – Parametric or Non-parametric

Examples of non-parametric distributions and impact on measures of central tendency

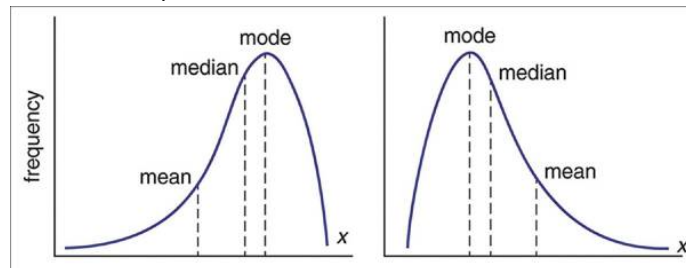


Image adapted under the Creative Commons license: Ali Z, Bhaskar SB. Basic statistical tools in research and data analysis. *Indian J Anesth.* 2016;60:662-669.

179

Statistics – Bayesian Analysis

- Statistical method that utilizes prior study information to analyze trial data where assumptions are not met in traditional statistical analyses
- For additional information:
 - Quintana M, et al. “Bayesian analysis: Using prior information to interpret the results of clinical trials.” *JAMA.* 2017;318(6):1605-1606.

180

Additional Sample Size Slides



181



Sample Size

Statistical Methods (Pg 6)

The sample size for this study was based on the assumption that 10% of the HIV-infected persons would be colonized during the study period [3,31,32] and that 85% in the placebo arm and 39% of the treatment arm would remain colonized based on prior literature [22], hence the sample size was estimated as 420 participants with a power 80% and alpha level of 0.05. Given potential loss to follow-up and that the colonization and clearance rates may vary, the study enrollment was *a priori* set at 550 participants. The sample size was also deemed adequate for the MRSA infection/SSTI outcome assuming 38% of the placebo group and 10% of the treatment group would develop infection based on a prior military study [13] and a subsequent study in HIV-infected persons [15], and accounting for the 6-month study visits over a two-year period.

Image adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

182

Sample Size

- To calculate sample size:
 - Study design (type, matching)
 - Decisional thresholds for alpha and beta
 - Anticipated variability for the primary outcome
 - Effect size
 - Anticipated size of the difference in the primary outcome between groups
 - Attrition (drop out)

183

Sample Size

Statistical Methods

The sample size for this study was based on the assumption that 10% of the HIV-infected persons would be colonized during the study period [3,31,32] and that 85% in the placebo arm and 39% of the treatment arm would remain colonized based on prior literature [22], hence the sample size was estimated as 420 participants with a power 80% and alpha level of 0.05. Given potential loss to follow-up and that the colonization and clearance rates may vary, the study enrollment was *a priori* set at 550 participants. The sample size was also deemed adequate for the MRSA infection/SSTI outcome assuming 38% of the placebo group and 10% of the treatment group would develop infection based on a prior military study [13] and a subsequent study in HIV-infected persons [15], and accounting for the 6-month study visits over a two-year period.

Primary Outcome Effect Size =
 $85\% - 39\% = 46\%$

30% drop out

Secondary outcome - accounts for
study design (repeated measures)

alpha=0.05

Beta=0.2

Image adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

184

Sample Size

- To calculate sample size:
 - Study design (type, matching); RCT (parallel)
 - Decisional thresholds for alpha and beta; $\alpha=0$; $\beta=0.2$
 - Anticipated variability for the primary outcome; ???
 - May have used variability in prevalence between studies (3 were cited)
 - Effect size; $85\% - 39\% = 46\%$
 - Attrition (drop out); 30% dropout

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

185

Sample Size

- Primary outcome:
 - Study states effectiveness
 - But sample size based on prevalence?
 - Ensures that sufficient numbers are tested to be randomized
- Our study enrolled 550 subjects
 - Met their sample size requirements?!?
 - But it is confusing, because only 49 were randomized
- <10% of the sample were MRSA positive
 - Assess for eligibility = 550; Not colonized=496; ~54 with MRSA

Adapted under the Creative Commons license. Weintrob A, Bebu I, Agan B, et al. Randomized, double-blind, placebo-controlled study on decolonization procedures for methicillin-resistant *Staphylococcus aureus* (MRSA) among HIV infected adults. *PLoS One*. 2015;10:e0128071.

186

Sample Size

- Sample size is affected by
 - Power
- The smaller the effect size, a larger the sample size is needed to detect a difference
- To decrease the likelihood of random error, increase the sample size
- With nonsignificant differences, consider adequacy of the sample size

187

Were washouts used? Were they of sufficient duration?

188

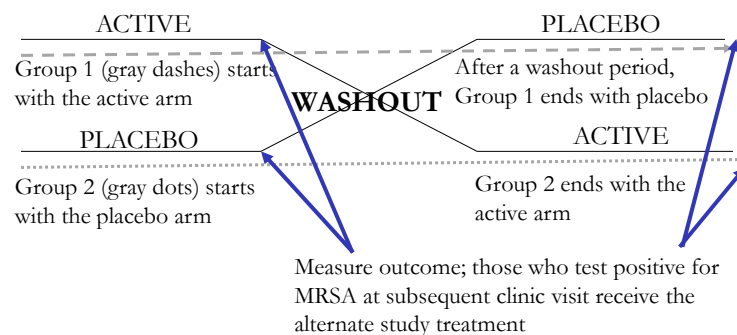
Washouts

- WHAT IF.....
 - The study was a cross-over study and patients that tested positive again received the alternative
 - Not typically done in ID studies

189

Washouts

- Two or more groups of patients
- Receive both placebo and intervention
- Washout period = 5 half-lives



190