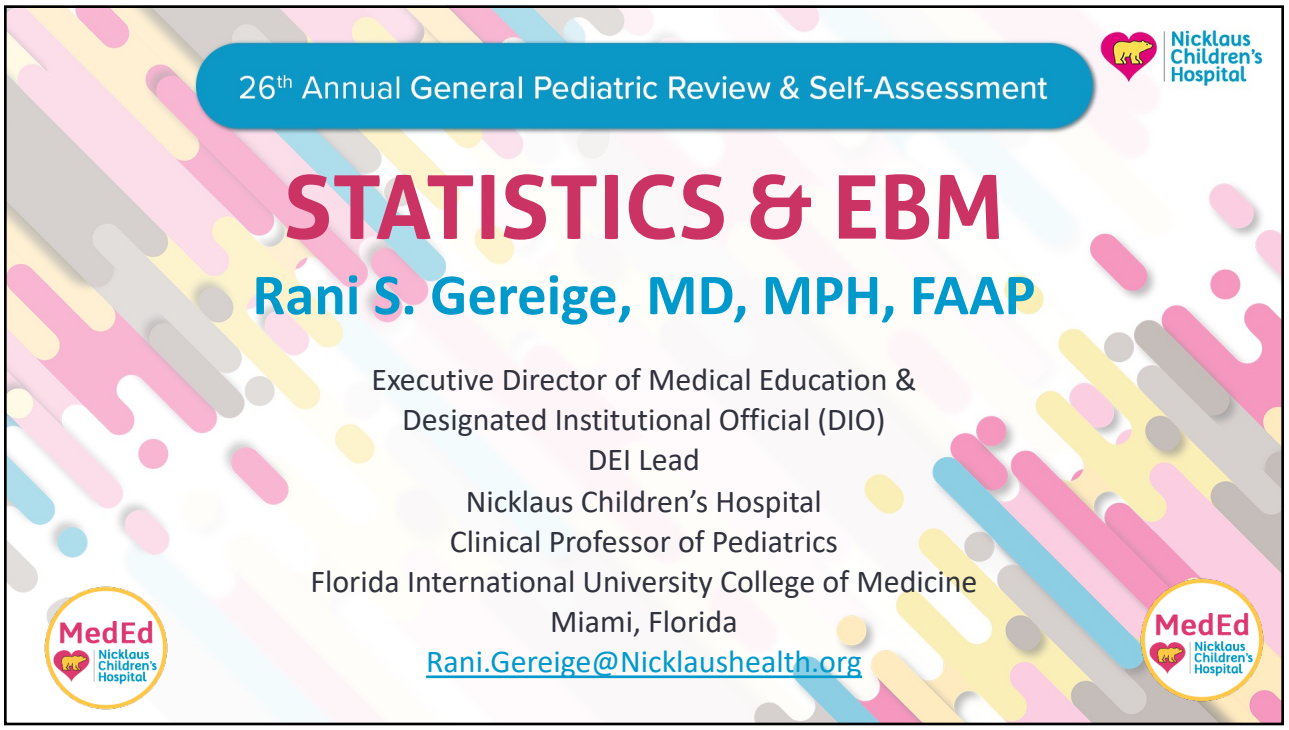
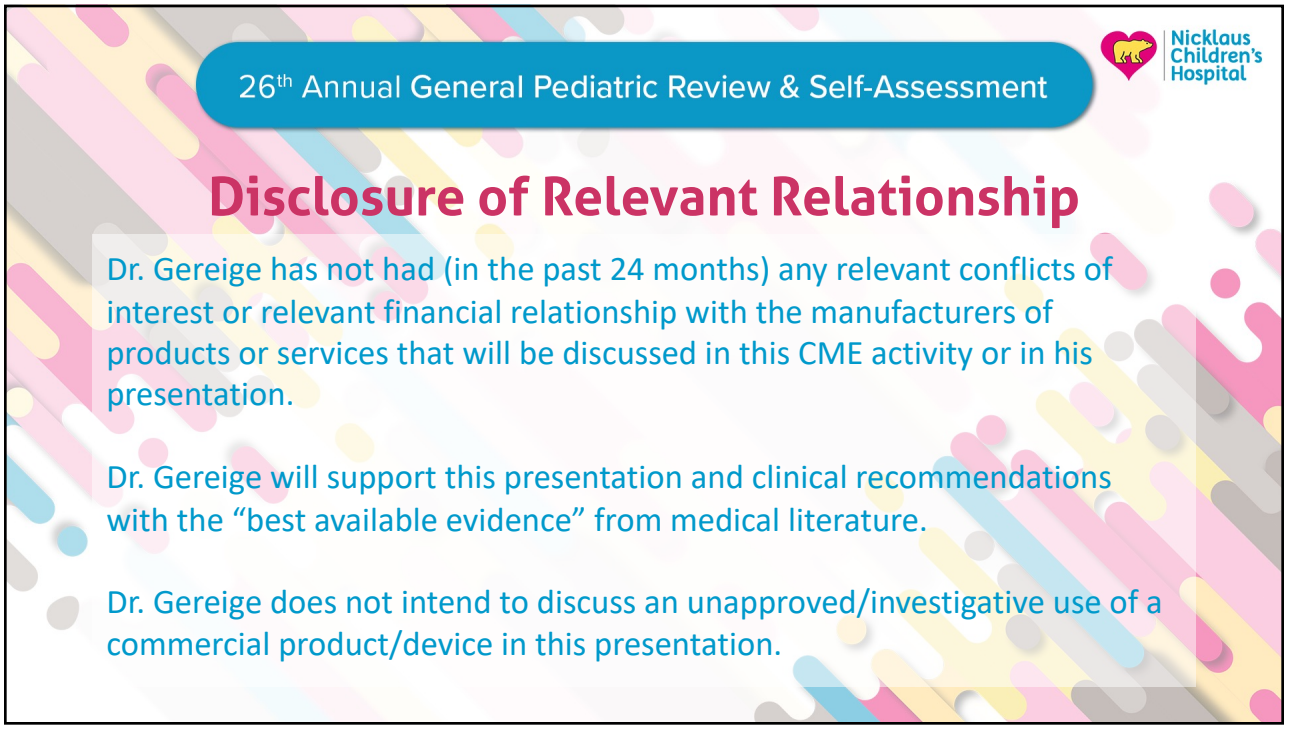

26th Annual General Pediatric Review & Self-Assessment
Nicklaus Children's Hospital
STATISTICS & EBM
Rani S. Gereige, MD, MPH, FAAP
Executive Director of Medical Education &
Designated Institutional Official (DIO)
DEI Lead
Nicklaus Children's Hospital
Clinical Professor of Pediatrics
Florida International University College of Medicine
Miami, Florida
Rani.Gereige@Nicklaushealth.org
MedEd
Nicklaus Children's Hospital
MedEd
Nicklaus Children's Hospital

1


26th Annual General Pediatric Review & Self-Assessment
Nicklaus Children's Hospital
Disclosure of Relevant Relationship
Dr. Gereige has not had (in the past 24 months) any relevant conflicts of interest or relevant financial relationship with the manufacturers of products or services that will be discussed in this CME activity or in his presentation.
Dr. Gereige will support this presentation and clinical recommendations with the "best available evidence" from medical literature.
Dr. Gereige does not intend to discuss an unapproved/investigative use of a commercial product/device in this presentation.

2

PREP Self-Assessment Content Specs (2013-2017)

What You Need to Know?

Nicklaus Children's Hospital

3

Research & Statistics

A. **Study design:** Understand the following:

1. **Validity hierarchy** for study design and study type
2. The uses and limitations of:
   - ✓ **Randomized clinical trials**
   - ✓ **Controlled clinical trials**
   - ✓ **Cohort studies**
   - ✓ **Case-control studies**
   - ✓ **Cross-sectional and longitudinal studies**
   - ✓ **Systematic review and meta-analysis**
   - ✓ **Descriptive epidemiologic studies**
   - ✓ **Case reports/series** and anecdotal evidence
3. How **sample size** affects the **power** of a study
4. How sample size may limit the ability to detect adverse events
5. Identify the **study design** most likely to yield valid information about:
   - ✓ The accuracy of a **diagnostic test**
   - ✓ The **benefits and/or harms** of an intervention
   - ✓ The **prognosis** of a condition

B. Data analysis

1. Understand:
   - ✓ **Validity** and how it might be compromised
   - ✓ **Reliability** and how it might be compromised
   - ✓ **Bias** and how it might distort the estimate of the association between exposure and outcome
   - ✓ **Confounding** and how to control for it in a study
   - ✓ **Generalizability** and how it relates to validity
   - ✓ The concept of **intention-to-treat analysis** to maintain the power of a study
   - ✓ The concept of **number-needed-to-treat** when utilized to describe therapeutic interventions
2. Distinguish between **type I and type II statistical errors**
3. Assess how the **data source** (eg, diaries, billing data, discharge diagnostic code) may affect study results

Ref: AAP - 2013-2017 PREP SA

Nicklaus Children's Hospital

4

Research & Statistics (Cont'ed)

C. Reading and interpreting results

1. Understand the following:
   - ✓ Prevalence and incidence
   - ✓ Pre-test and post-test probability
   - ✓ Positive and negative predictive values
   - ✓ Sensitivity and specificity and how to apply them to test results
   - ✓ Standard deviation in the interpretation of results
   - ✓ Standard error in the interpretation of results
   - ✓ Confidence interval in the interpretation of results
   - ✓ Likelihood ratio and when it might be useful to reach a diagnosis
   - ✓ Relative risk analysis and odds ratio
2. Distinguish statistical significance from clinical importance
3. Given the need for specific clinical information, identify a clear, structured, searchable clinical question

Ref: AAP - 2013-2017 PREP SA

Nicklaus Children's Hospital

5

Outline for Today's Lecture

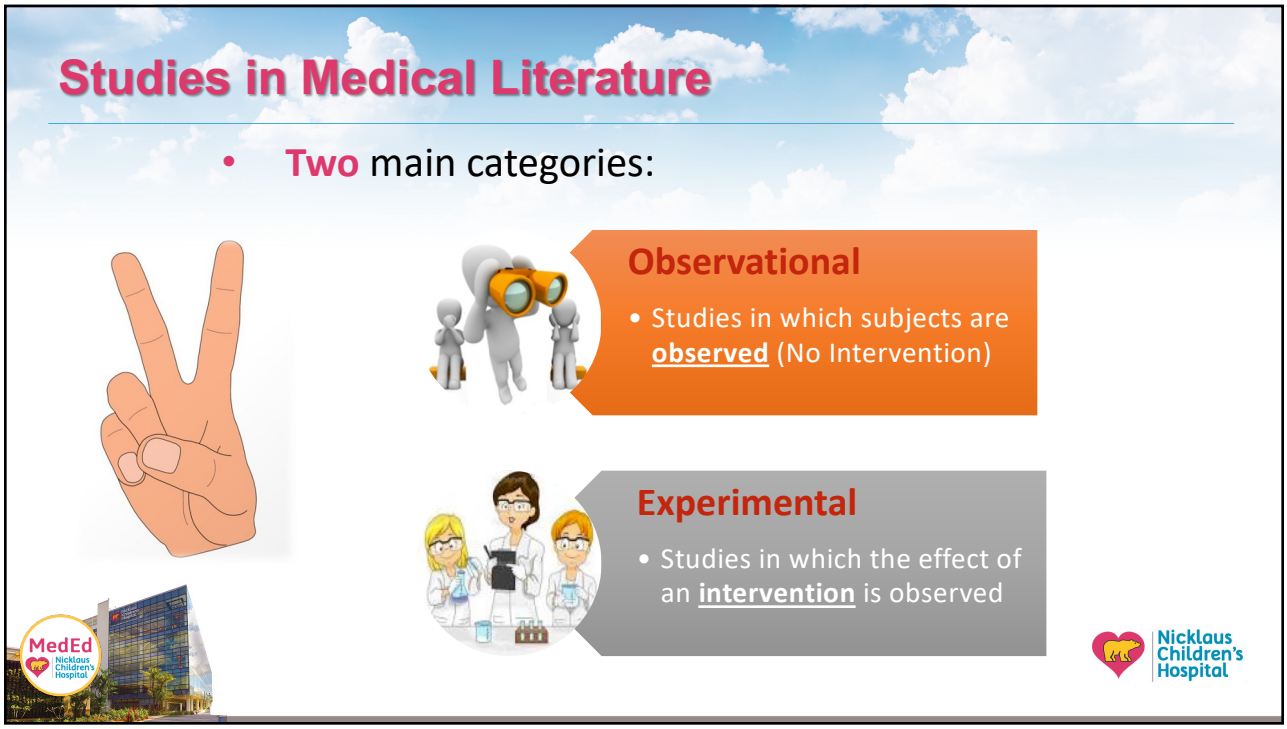
- **Research/ Study Design**
  - Types of Studies: Observational Vs Experimental (Strengths & Limitations)
  - Hierarchy of Research Design
- **Data Analysis**
  - Accuracy (Validity) & Precision (Reliability)
  - Bias & Confounding: Ways to minimize Bias and Confounding
  - Relative Risk (RR); Absolute Risk Reduction (ARR); Relative Risk Reduction (RRR); & Number Needed to Treat (NNT)
  - Types of Data
  - Type I & Type II Errors
  - P-Value, Power, Effect Size
- **Reading & Interpreting Results (EBM)**
  - Clinical Questions
  - Therapy Studies (NNT), Intention-to-Treat
  - Harm/Association (NNH)
  - Diagnostic Tests (Sens, Spec, PPV, NPV, LRs, Pre & Post-Test Probability)
  - Prognosis

Nicklaus Children's Hospital

6


Research & Study Design
Result 3
Result 2
Result 1
Survey:
Excellent:
Good:
Fair:
Poor:
MedEd
Nicklaus Children's Hospital

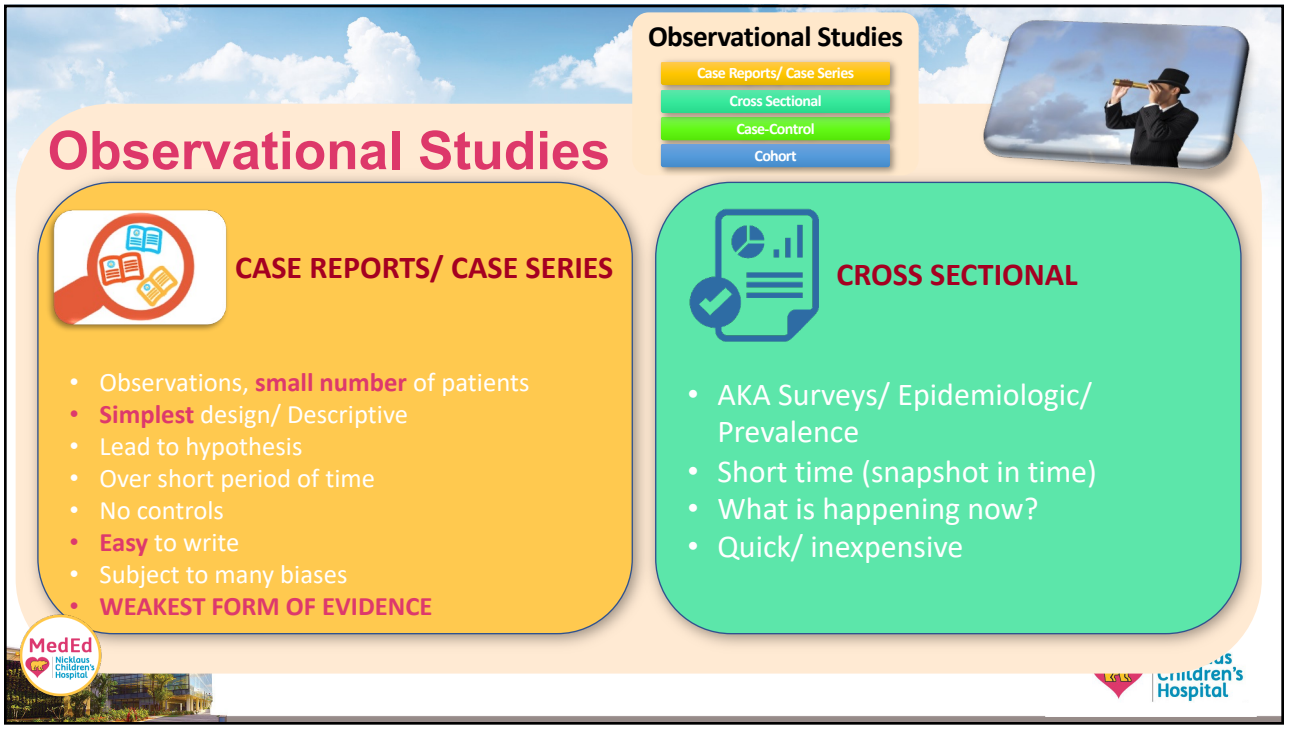
7


Studies in Medical Literature
• Two main categories:
Observational
• Studies in which subjects are observed (No Intervention)
Experimental
• Studies in which the effect of an intervention is observed
MedEd
Nicklaus Children's Hospital

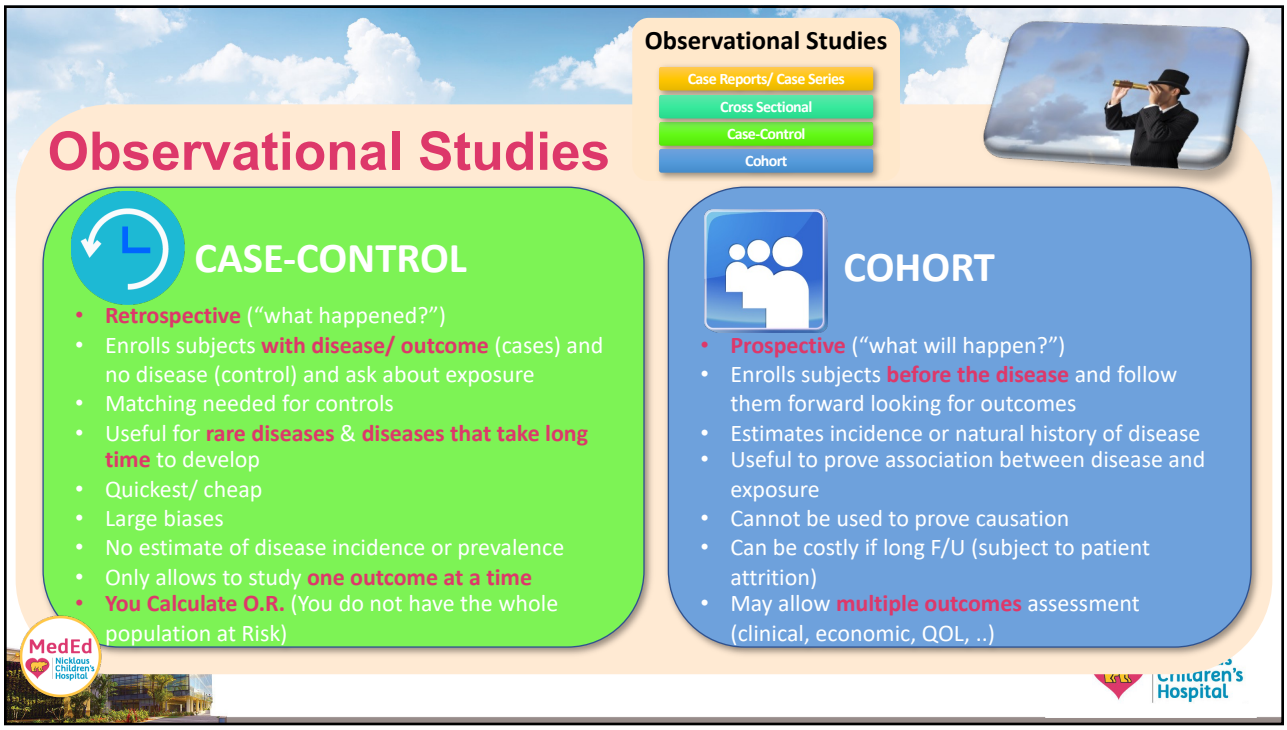
8


Observational Studies
Case Reports/ Case Series
Cross Sectional
Case-Control
Cohort
MedEd
Nicklaus Children's Hospital

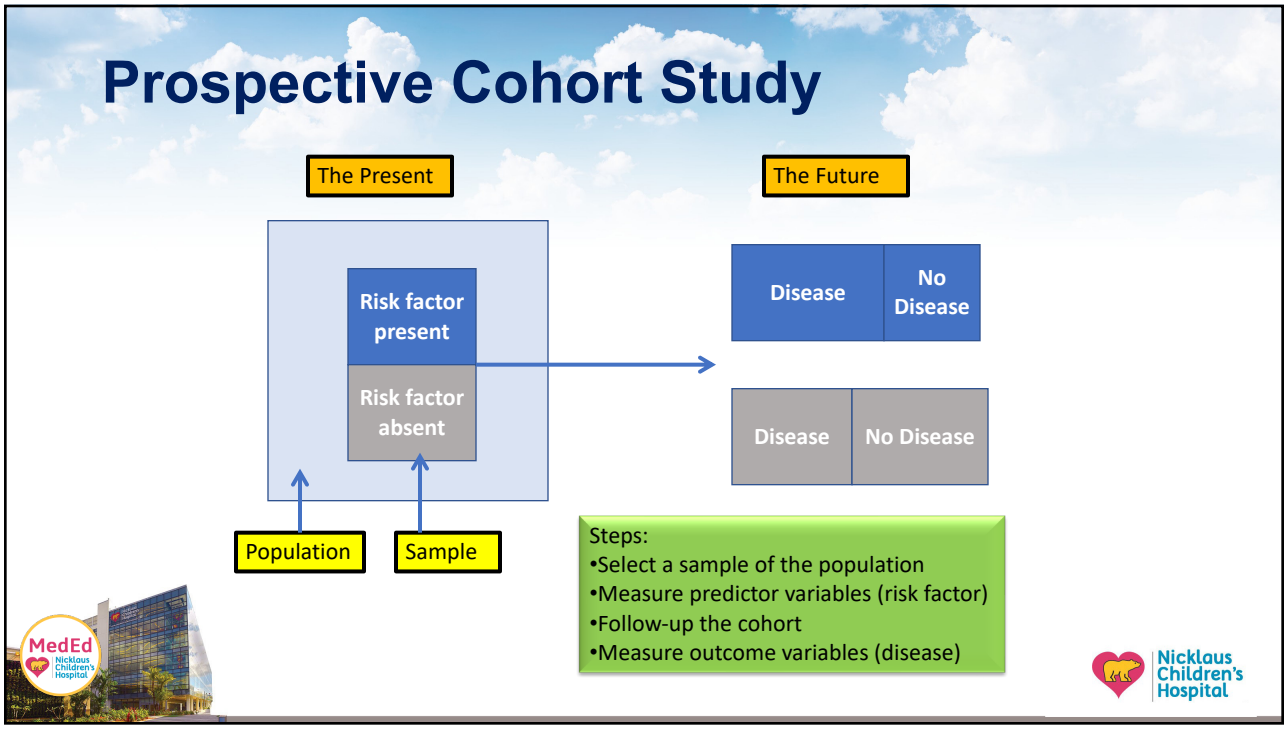
9


Observational Studies
Case Reports/ Case Series
Cross Sectional
Case-Control
Cohort
Observational Studies
CASE REPORTS/ CASE SERIES
Observations, small number of patients
Simplest design/ Descriptive
Lead to hypothesis
Over short period of time
No controls
Easy to write
Subject to many biases
WEAKEST FORM OF EVIDENCE
CROSS SECTIONAL
AKA Surveys/ Epidemiologic/ Prevalence
Short time (snapshot in time)
What is happening now?
Quick/ inexpensive
MedEd
Nicklaus Children's Hospital

10


Observational Studies
Case Reports/ Case Series
Cross Sectional
Case-Control
Cohort
Observational Studies
CASE-CONTROL
Retrospective ("what happened?")
Enrolls subjects with disease/ outcome (cases) and no disease (control) and ask about exposure
Matching needed for controls
Useful for rare diseases & diseases that take long time to develop
Quickest/ cheap
Large biases
No estimate of disease incidence or prevalence
Only allows to study one outcome at a time
You Calculate O.R. (You do not have the whole population at Risk)
COHORT
Prospective ("what will happen?")
Enrolls subjects before the disease and follow them forward looking for outcomes
Estimates incidence or natural history of disease
Useful to prove association between disease and exposure
Cannot be used to prove causation
Can be costly if long F/U (subject to patient attrition)
May allow multiple outcomes assessment (clinical, economic, QOL, ..)
MedEd
Nicklaus Children's Hospital

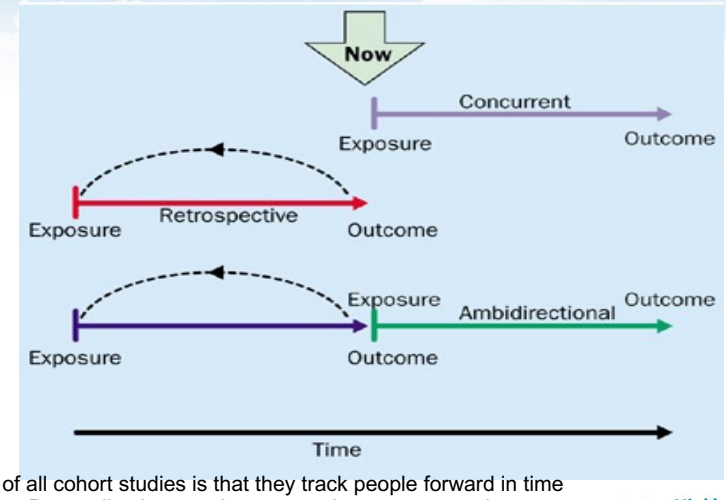
11


Prospective Cohort Study
The Present
The Future
Risk factor present
Risk factor absent
Disease
No Disease
Disease
No Disease
Population
Sample
Steps:
•Select a sample of the population
•Measure predictor variables (risk factor)
•Follow-up the cohort
•Measure outcome variables (disease)
MedEd
Nicklaus Children's Hospital

12

Cohort studies: marching towards outcomes (Prospective)

Looking forward in the past = Retrospective Cohort

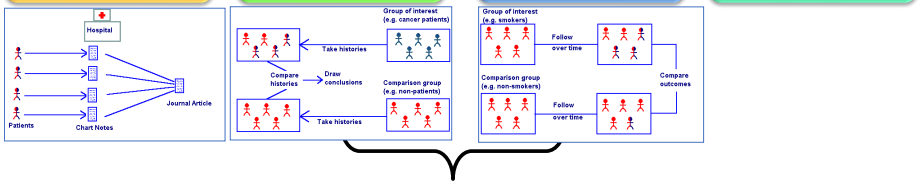


Lancet 2002; **359:** 341-45

The defining characteristic of all cohort studies is that they track people forward in time **from exposure to outcome**. Data collection may be prospective or retrospective. **Ex. Contraceptives and DVT**.

13

Observational Studies



Longitudinal Studies

"Notion of Time"

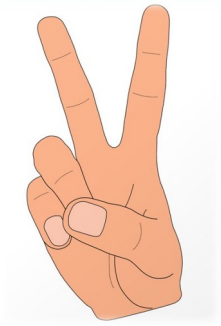
Table. **Summary of Designs**

Study Design	Definition	Strengths	Weaknesses
Cross-Sectional	Single data collection point	• Quick • Inexpensive • Establishes prevalence • Suggests future research directions	• Difficult to determine causality • Possible spurious associations
Longitudinal Prospective Retrospective	Multiple data collection points occur over time	• Can determine causality • Can monitor trends • Less concerned with spuriousness	• Time-consuming • Expensive

14

Studies in Medical Literature

- **Two** main categories:



Observational

- Studies in which subjects are **observed** (No Intervention)

Experimental

- Studies in which the effect of an **intervention** is observed

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

15

Experimental Studies

AKA "Clinical Trials"

(easy to identify, explicitly stated in the abstract, Expensive)

Experimental Studies

Controlled Trials

Uncontrolled Trials

Self-Controls

- Subject to bias (Hawthorne effect)
- Can do crossover study (with washout period in between)

Independent Concurrent Controls

External Controls

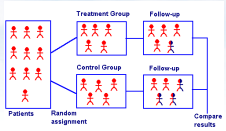
- Uses the results of another investigator's research as a comparison
- Historical controls can also be used: for disease with no cures yet

RCT

- Considered the "gold standard"
- Double or single blind
- **The epitome of all research designs**
- Provides the strongest evidence of concluding causation
- Best insurance that results are due to the intervention

Non-Randomized

- Opened to biases

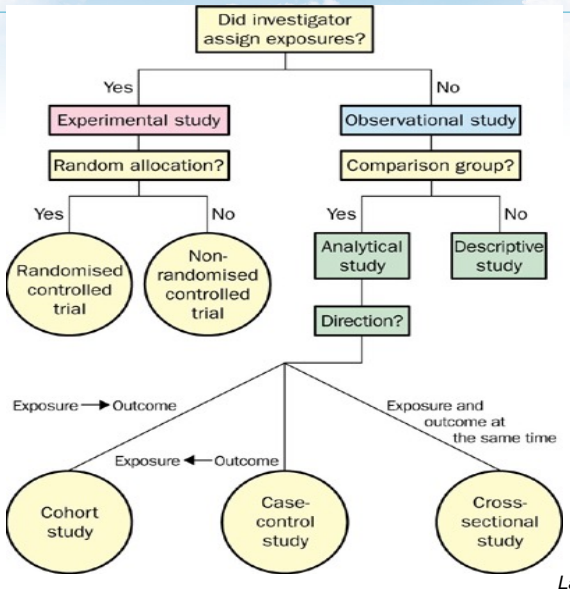


MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

16

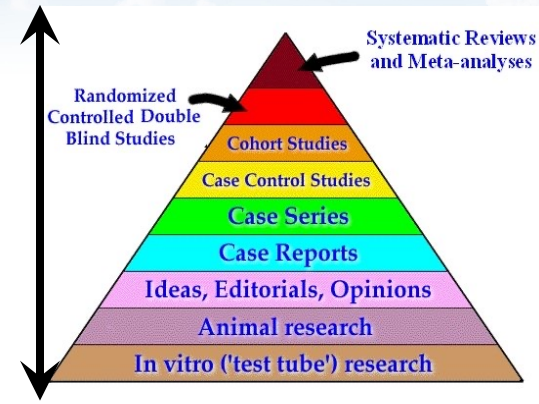
Classification of Types of Clinical Research



Lancet 2002; **359:** 57-61

17

Hierarchy of Research Design



Advantages and Limitations

- Pool results from **multiple** studies
- Findings offer a compilation of evidence (Greater power than an individual study)
- **Meta-analysis of multiple RCTs is the best**
- Meta-analysis = Systematic Review + analysis of results of multiple studies.
- The most significant limitation of both systematic reviews and meta-analyses is commonly described as "garbage in, garbage out."

18

Data Analysis

19

Validity

EBM/ Statistics

Internal Validity

- Accuracy of study's conclusion
- Needed to determine the causal relations among variables

Clinical Application

External validity

- How well the study represents the "real world"
- Generalizability

20

Accuracy and Precision

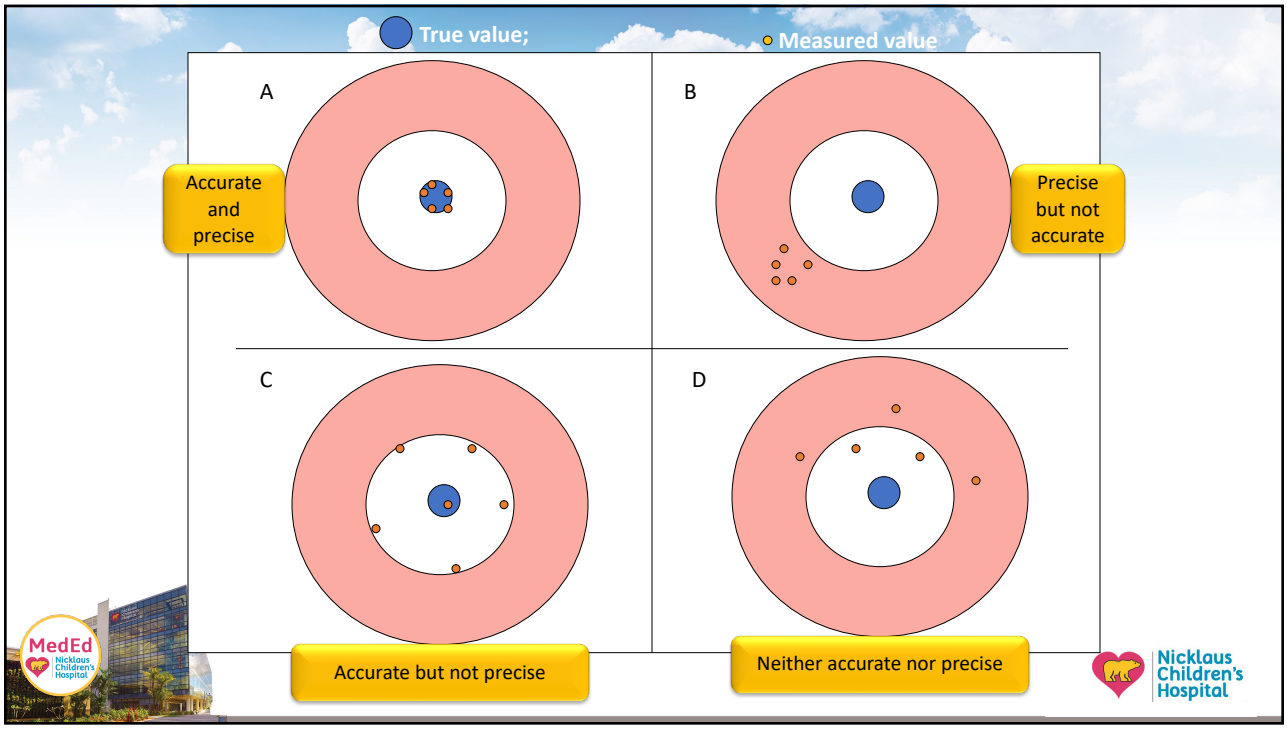
Accuracy (Validity)

- The closeness of a measurement to the true value of the quantity that is measured.
- Affected by **systematic errors**

Precision (Repeatability; reliability)

- The closeness of agreement of two or more measurements of the same quantity.
- Affected by **random errors**

21



22

Table. Validity Hierarchy

Internal Validity

Study Design	Strengths	Weaknesses
Randomized controlled trials	• High internal validity • Reduced risk of confounding variables	• Reduced external validity • Expensive, time-consuming
Cohort studies	• Useful for sequential events • Can study multiple outcomes • *Retrospective:* less expensive	• Requires large sample size • Risk of confounding variables • Difficult to study rare outcomes • *Prospective:* Expensive
Case-control studies	• Useful for rare outcomes • Can study several exposures • Inexpensive	• Risk of confounding variables
Cross-sectional studies	• Can study multiple outcomes and exposures	• Cannot infer causality • Risk of confounding variables • Less useful for rare exposures or outcomes
Case studies	• Useful for rare outcomes • Convenient, inexpensive	• Risk of confounding variables • Lack of a comparison group • Cannot infer causality

ted from Ho, et al. *Circulation*. 2008;118:1675–1684.

23

Reliability

Test-retest reliability:

- Assesses whether an instrument or test yields the **same results each time it is used** with the same study sample under the same study conditions
- One way to determine whether an instrument or test is reliable or consistent is to administer it with the **same subject or sample more than once**

Internal consistency reliability:

- A measure of the consistency of the items within a test.

Inter-rater reliability:

- The degree to which two raters independently score an observation similarly.

Nicklaus Children's Hospital

24

Study of Any Intervention

- *Random Error* = Deviation from the underlying truth by chance
- *Bias* = Systematic deviation from underlying truth
  - Definition of Bias:
    - "Any systematic error in the design, conduct or analysis of a study that results in a mistaken estimate of an exposure's effect on the risk of disease."

Nicklaus Children's Hospital

25

Bias

- Major issue in epidemiologic research studies
- Can lead to inferences that systematically deviate from truth

Nicklaus Children's Hospital

26

Common Types of Bias

- *Surveillance bias:* Population being monitored more closely or more frequently than the general population
- *Selection bias:* Two primary varieties:
  1. Systematic differences in the characteristics between individuals selected for a study compared with those not selected for the study
  2. Systematic differences in the selection of cases and controls or exposed and unexposed individuals
- *Misclassification bias: M*isclassifying individuals into diseased or non-diseased groups or into exposed and unexposed groups

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

27

Ways to Decrease Bias

(During Design, Before Completion, and After Completion)

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

28

(During Selection of Control Group) Minimizing Selection Bias

STRATEGY	COMMENT
Restriction or Specification	Limits the range of characteristics of the patients in the study, decreases sample size, heterogeneity and generalizability (External Validity)
Matching	For each patient in the study group, select one or more patients with the same characteristics for a comparison group
Adjustment	Mathematical corrections to create an equal weight for dissimilar characteristics
Stratification	Compare outcomes from subgroups of each group with similar characteristics (i.e. age by decades)
Randomization	Randomization of the study population and controls

Ref: Clinical Epidemiology The Essentials. 3rd Ed. Fletcher et al. 1996, p 129.

29

(Before Study Completion) Ways to Decrease Bias

STRATEGY	BENEFITS	TRADE-OFF
Limitations for participation (Exclusion Criteria)	•By restricting the heterogeneity of the group, we reduce the opportunity for differences in outcome that aren't due to the treatment itself •**Improves INTERNAL VALIDITY**	Makes generalization of the results more precise but **limits EXTERNAL VALIDITY/ GENERALIZABILITY** to a smaller portion of the population
Use of a Control/ Comparison Group	Minimizes the 'Hawthorne effect" *By virtue of being in a study, the patient's behavior changes and has a better prognosis*	Still may have a "placebo effect" unless placebo given to control group *Giving a pill with an expected/potential result can provide effect even if the pill is inert*

30

Confounding

- One of several threats to internal validity of a research study
- Confounding is defined as:
  - A possible source of bias in studies in which an unmeasured third variable (the confounder) is related to the exposure of interest (although not causally) and causally related to the outcome of interest

Nicklaus Children's Hospital

31

(After Study Completion) Dealing With Confounding Two ways

Stratification

- Subdividing subjects by levels of a potential KNOWN confounding variable
- Testing for the association of exposure with outcome within each stratum
- **Disadvantages:**
  - May not be feasible to handle multiple confounders
  - As the number of strata increase, sample size within each stratum decreases, reducing statistical power
  - May not adequately control for confounding

Multivariate Techniques

- Permit understanding of how much variability in an outcome is accounted for by a confounder
- Permit researchers to control for more factors than stratification
- **Disadvantages:**
  - Require readers to understand how to interpret the meaning of adjusted odds ratios and regression coefficients as well as how statistical significance was determined

Nicklaus Children's Hospital

32

Randomization

What?

- Participation in a study arm by **chance**, not by choice
- Equal & fair chance of getting intervention or control

Goals

- **Produce comparable groups** in terms of general participant characteristics (known and unknown confounders)
- The two groups will be similar at the baseline
- **Avoids selection bias**

How?

- **Simple randomization**: repeated fair coin-tossing; good for large sample
- **Block randomization**: subjects randomized in a block (of 6 or 8) to prevent uneven allotment in a small sample
- **Stratified randomization**: randomize to groups according to covariates (like age groups under and over 12 years)

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

33

Blinding & Allocation Concealment

Blinding/ Masking

- **Types of blinding**
  - Single-blind (subject or care giver)
  - Double-blind (subject and care giver)
  - Triple-blind (subject, care giver and data analyzer)
- **Open trial:** unblind or unmasked study
- **Benefit:** Removes the bias
  - Placebo effect
  - Observer bias (change in behavior due to the awareness of being observed; Hawthorne Effect)
  - Experimenter bias

Allocation Concealment

- Not same as "blinding"
- The subject and the investigator do not know the allocation of the group until randomization
  - After allocation it may or may not be blinded

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

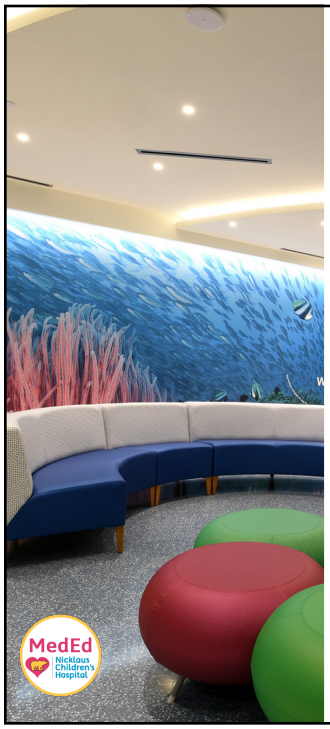
34

Potential Bias	Strategy Against
Sampling Bias	• Target population
Selection Bias	• Randomization
Placebo Effect	• Placebos • Blinding participants
Cointerventions	• Blinding Providers • Treatment protocols
Assessment Bias	• Blinding Providers
Follow-up	• Ensuring completeness

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

35



MedEd Nicklaus Children's Hospital

Relative Risk (RR) vs Odds Ratio (OR)

Nicklaus Children's Hospital

36

What is Risk?

- Risk is the probability of an event!
- In statistical terms: Risk is P(e)

In statistics, RISK doesn't mean harmful events!

Nicklaus Children's Hospital

37

Probability vs Odds

- P = Odds/(1+Odds)
- Odds = P/(1-P)

PROBABILITY	ODDS
P(e) = e/n = events/ (events + non-events)	Odds = events/ non-events
P = # of events/ Total possible events	
EXAMPLE	**EXAMPLE**
In deck of cards: P(spade)	*In deck of cards: Odds(spade)*
• *13 events (spades in the deck)*	• *13 events (spades in the deck)*
• *39 non-events (52-13, non-spade cards)*	• *39 non-events (52-13, non-spade cards)*
• *P(spade) = 13/ (13+39) = 13/52 = 1/4*	• *Odds for spade = 13/39 = 1/3*
Probability of picking a spade from a deck of cards is 1 in 4	*The odds of picking a spade from a deck of cards is 1 in 3*

Pearl to Remember

Probability is always smaller than odds. P vs O

Hence, the denominator is larger to calculate probability from odds (1+odds) compared to odds from probability (1-probability)

38

Odds Ratio (OR)

- Ratio of Odds
- The odds of a case patient being exposed divided by the odds of a control patient being exposed
- Calculated in Case-Control studies
- Proportion exposed in a diseased vs. non-diseased patient sample
- OR > 1 represents an increased risk or association
- Describes the relative harm of an exposure independent of disease prevalence
- When the prevalence of the outcome of interest is rare in the population from which the sample was drawn (often the reason for using a case-control study), the OR closely approximates the RR

Nicklaus Children's Hospital

39

Relative Risk (RR); Absolute Risk Reduction (ARR); Relative Risk Reduction (RRR); & Number Needed to Treat (NNT)

- In an experiment the mortality fell to 30% with a new treatment [P(E)] from 40% in control group [P(C)].
  - What is the relative risk (RR) of mortality with the new treatment?
  - What is the absolute risk reduction (ARR) of mortality with the new treatment?
  - What is the relative risk reduction (RRR) of mortality with the new treatment?
  - How many patients need to be treated with the new medication to avoid one death?

Nicklaus Children's Hospital

40

Relative Risk (RR)

- **Relative risk** = the ratio of the risk of an event in experimental group to the risk of the same event in control (or other) group.

$$RR = P(E)/ P(C)$$

- From the previous example:
  - P(E) = 0.3 (30%)
  - P(C) = 0.4 (40%)
  - RR = P(E)/P(C) = 0.3/0.4 = 0.75 (75%)
  - The relative risk of mortality with the new treatment is 0.75
    - The mortality in the Rx group is 75% of that in control group

41

Absolute Risk Reduction (ARR)

- **Absolute risk reduction (ARR)** or called **risk decreased (RD)** = the difference in the risks of an event in two groups.

$$ARR = P(C) – P(E)$$

- In the example given:
  - P(E) = 0.3
  - P(C) = 0.4
  - ARR = (0.4 – 0.3) = 0.1 (10%)
  - Mortality risk is 10% less with the new Rx compared to that in control group

42

Relative Risk Reduction (RRR)

- **Relative risk reduction (RRR)** = the ratio of absolute risk reduction (ARR) to the risk in control group.

$$RRR = (P(C) - P(E))/P(C) = (ARR)/P(C)$$

- In the example given:
  - P(E) = 0.3
  - P(C) = 0.4
  - RRR = (0.4 – 0.3)/0.4 = 0.1/0.4 = 0.25 (25%)
  - Mortality risk is reduced by 25% with the new Rx compared to that in control group

43

Number Needed to Treat (NNT)

$$NNT = 1/ARR$$

- In the example given:
  - P(E) = 0.3
  - P(C) = 0.4
  - ARR = (0.4 – 0.3) = 0.1
  - NNT = 1/ARR = 1/0.1 = 10
  - We need to treat 10 patients with the new Rx to avoid one death

44

Intention-to-Treat Principle

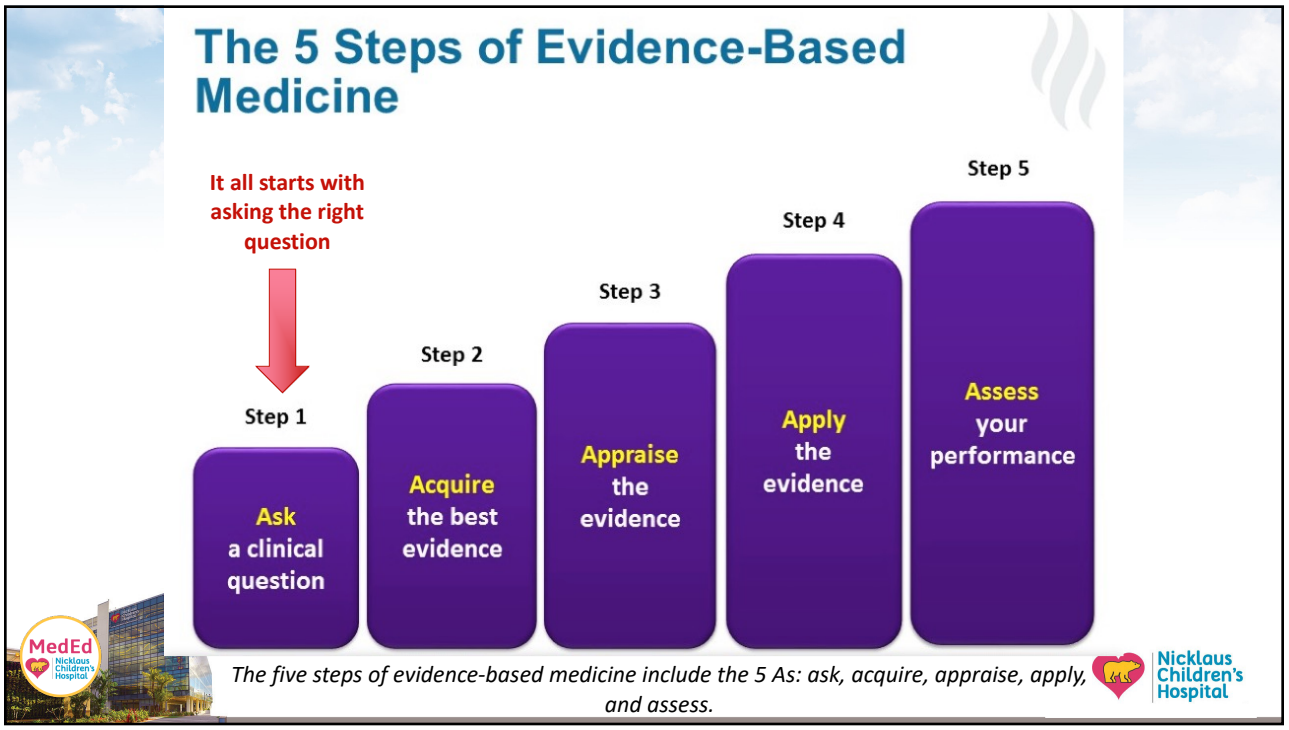
- **GOAL:** Preserves the randomization of unknown confounders
- Include all patients in the group they have randomized to, irrespective of the treatment received or not
- Include the subjects in the original group for analysis even if
  - They have stopped receiving the study intervention
  - They have crossed over to the counter intervention
  - They were lost to follow-up
  - Died
  - Left the study

45

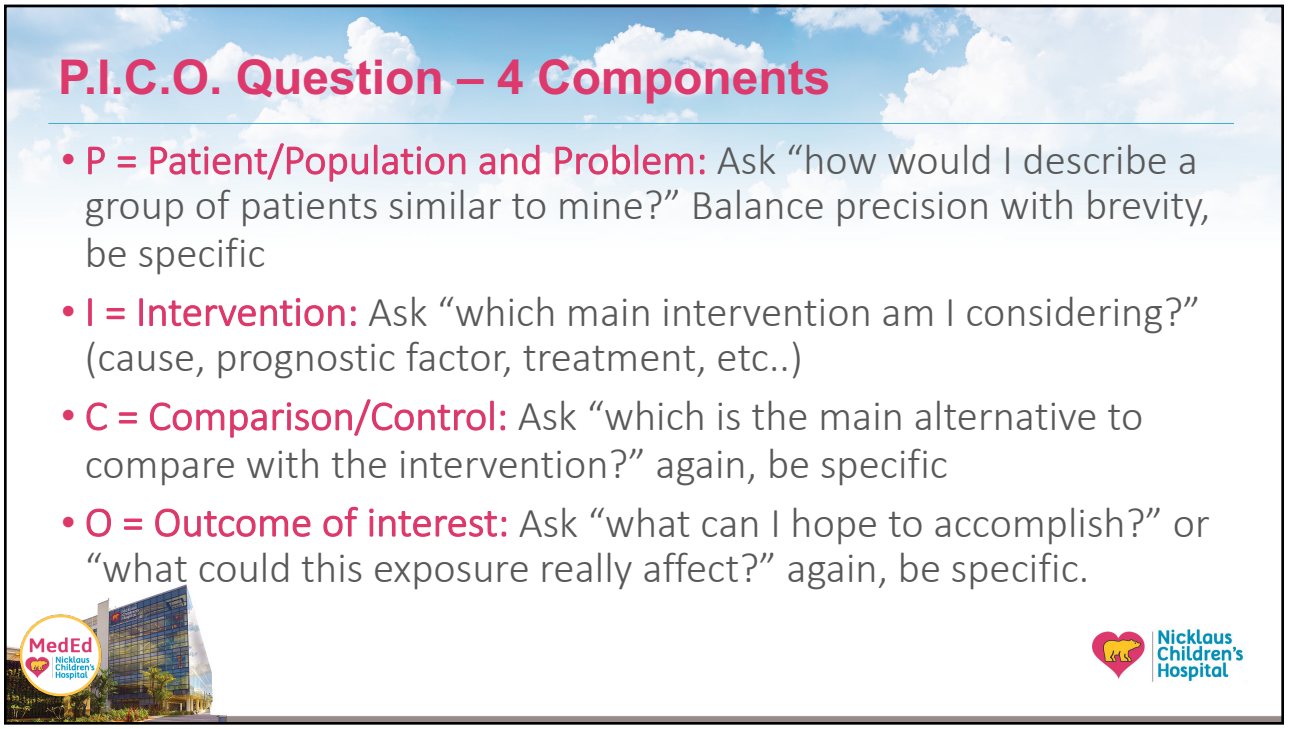
Reading & Interpreting Results (EBM)

Nicklaus Children's Hospital

46


The 5 Steps of Evidence-Based Medicine
It all starts with asking the right question
Step 1
Ask a clinical question
Step 2
Acquire the best evidence
Step 3
Appraise the evidence
Step 4
Apply the evidence
Step 5
Assess your performance
MedEd
Nicklaus Children's Hospital
The five steps of evidence-based medicine include the 5 As: ask, acquire, appraise, apply, and assess.
Nicklaus Children's Hospital

47


P.I.C.O. Question – 4 Components
• P = Patient/Population and Problem: Ask "how would I describe a group of patients similar to mine?" Balance precision with brevity, be specific
• I = Intervention: Ask "which main intervention am I considering?" (cause, prognostic factor, treatment, etc..)
• C = Comparison/Control: Ask "which is the main alternative to compare with the intervention?" again, be specific
• O = Outcome of interest: Ask "what can I hope to accomplish?" or "what could this exposure really affect?" again, be specific.
MedEd
Nicklaus Children's Hospital
Nicklaus Children's Hospital

48

PICO Question

Table 1. Example of the PICO Process to Identify a Research Question

P	Patient or Population	Newborns
I	Intervention or Indicator	Newborn screen for congenital hypothyroidism
C	Comparison	Serum TSH measurement
O	Outcome	Diagnosis of congenital hypothyroidism or the "accuracy of the test"

Table 3. Sample Search Terms

	Key Concepts	Search Term
P	Newborns	Newborn or infant
I	Newborn screen for congenital hypothyroidism	Newborn screen for congenital hypothyroidism or congenital hypothyroidism diagnosis
C	Compared with TSH	
O	Accuracy of test	Accuracy or diagnosis

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

49

(5) Incidence vs Prevalence

Incidence

- Quantifies the # of **new events** or cases of disease that develop in a population of individuals at risk during a **specified time interval**
- **# new cases** of a disease during a given period of time /total population at risk
- Provides an estimate of the probability (risk) that an individual will develop a disease during a specified period of time

Prevalence

- Quantifies the **proportion of individuals in a population** who have the disease at a specific instant
- Provides an estimate of the probability (risk) that an individual will be ill at a point in time
- **# of existing cases** / the total population (at a given point in time)

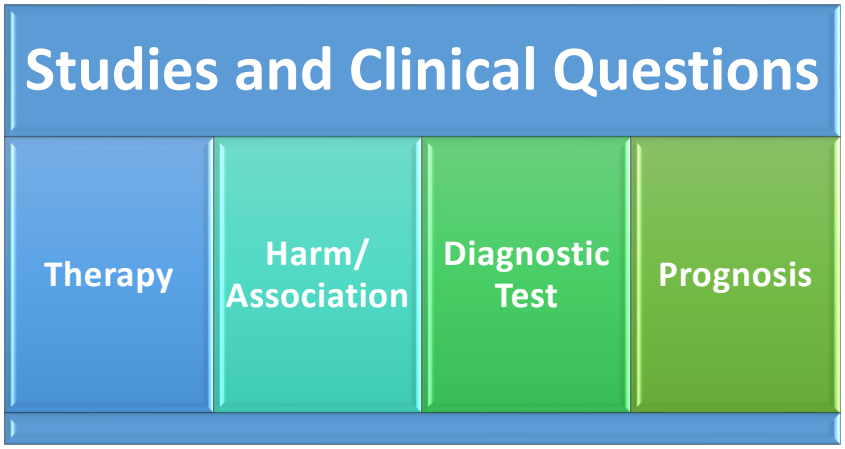
MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

50

Studies and Clinical Questions

- Four possible Domains:



51

Therapy Studies

- Best is RCT, followed by Cohort
- Subjects randomized to:
  - New Treatment (A) or Old Treatment (B)
  - New Treatment or Placebo
- Outcome is measured (Improvement)

You Calculate a NNT

NNT = 1/ARR

52

Harm/ Association Studies

- RCT unethical
- Cohort is next best
- Outcome is measured (Harm)

You calculate a NNH

NNH = 1/ARI

Nicklaus Children's Hospital

53

Case-Control Studies
Odds Ratio (Relative Odds)

Adverse Outcome

	Present	Absent	Totals
Exposure Yes	a	b	a+b
Exposure No	c	d	c+d
Totals	a+c	b+d	a+b+c+d

OR = (a/c)/(b/d) = ad/bc

Nicklaus Children's Hospital

54

Association is Different from Causation

- Five criteria must be fulfilled to prove causation:

1. Is it clear that the exposure preceded the onset of the outcome? – Looks at exclusion criteria
2. Is there a dose-response gradient?
   - e.g. Smoking and lung cancer
3. Is there any positive evidence from a de-challenge / re-challenge study?
4. Is the association consistent from study to study?
5. Does the association make biological sense? - Pathophysiology

Nicklaus Children's Hospital

55

Diagnostic Test Studies

- All subjects receive the new test and the "gold standard"
  e.g. Rapid Strep and throat culture

- 4 possibilities

By Gold Standard

	Patients with disease	Patients without disease
Test is positive	a true positive	b false positive
Test is negative	c false negative	d true negative

Nicklaus Children's Hospital

56

Evaluating the Evidence – Diagnostic Test

Construct the 2x2 table

	Patients with disease	Patients without disease
Test is positive	a true positive	b false positive
Test is negative	c false negative	d true negative

Sensitivity = a /a+c = P(+/D)= P(TP among diseased) = TP/(TP+FN)
Specificity = d /b+d = P(-/Ø)= P(TN among non diseased) = TN/ (FP+TN)

PPV = a /a+b = P(D/+)= P(TP among all Positives) = TP/ (TP+FP)
NPV= d /c+d = P(Ø/-)= P(TN among all Negatives) = TN/ (FN+TN)

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

57

Sensitivity & Specificity of a Test

Sensitivity

- Ability of a test to recognize correctly persons who ***have a disease*** or condition
- Proportion of patients who have a disorder in whom the results of the test are *positive*

PID

(Positive in Disease)

Specificity

- Ability of a test to recognize correctly persons who ***do not have a disease*** *or* condition
- Proportion of patients who do not have a disorder in whom the test result is *negative*

NIH

(Negative in Health)

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

58

SpPin & SnNout

- SpPin = Result of a test with high Specificity, when Positive, rules in the diagnosis
- SnNout = Result of a test with high Sensitivity, when Negative, rules out the diagnosis

Discriminant ability of a test = (sensitivity+specificity)/2

59

Predictive Values of a Test

PPV	NPV
• Proportion of patients ***testing positive*** who actually have the disease or condition in question	• Proportion of patients ***testing negative*** who actually *do not* have the condition in question

60

IMPORTANT

- Sensitivity and specificity are properties intrinsic to a test and are not affected by the prevalence of a particular disease or condition
- The predictive values of a diagnostic test are influenced greatly by prevalence. The higher the disease prevalence, the higher the PPV. e.g. The rapid flu test has a higher PPV during the flu season (time of high prevalence)

61

Pretest Probability → Post-test Probability
Baye's Theorem

LR calculator: http://getthediagnosis.org/calculator.htm

- Varies by: Physician experience, season, geography, prevalence, and the history and physical findings
- Clinician's best estimate of the probability of a specific disease **before** diagnostic testing
- Generally has a large impact on the diagnostic process

Pre-test Probability (X)

Likelihood Ratio
LR(+) or LR(-)

- Links Sensitivity and Specificity
- Indicates by how much a given diagnostic test result will raise or lower the pretest probability of the target disorder

Post-test Odds: Z = Y x LR

Pre-test Odds: Y = X/(100-X)

Post-test probability = Z/(Z+1)

- Depends on the magnitude of LR
- Is the clinician's best estimate of the probability of a specific disease **after** diagnostic testing

Post-Test Probability

62

Likelihood Ratios

LR(+)

- Probability of person WITH disease having positive test/probability of person WITHOUT disease having a positive test
- P(TP)/P(FP)
- LR(+) = Sens/(1-spec)
- Corresponds to clinically "ruling in disease"

LR(-)

- Probability of person WITH disease having negative test/probability of person WITHOUT disease having negative test
- P(FN)/P(TN)
- LR(-) = (1-sens)/spec
- Corresponds to clinically "ruling out disease"

Indicate by how much a given diagnostic test result will raise or lower the pretest probability of the target disorder

- LR = 1 → Post-test probability = Pre-test probability
- LR > 1 → increases the probability that the target disorder is present
- LR < 1 → decreases the probability that the target disorder is present

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

63

Guide to the Significance of LRs

LR	Interpretation
> 10	Large and often conclusive increase in the likelihood of disease
5 - 10	Moderate increase in the likelihood of disease
2 - 5	Small increase in the likelihood of disease
1 - 2	Minimal increase in the likelihood of disease
1	No change in the likelihood of disease
0.5 - 1.0	Minimal decrease in the likelihood of disease
0.2 - 0.5	Small decrease in the likelihood of disease
0.1 - 0.2	Moderate decrease in the likelihood of disease
< 0.1	Large and often conclusive decrease in the likelihood of disease

- LR > 10 or < 0.1 generate large and often conclusive changes from pre-test to post-test probability
- LR = 5 - 10 or 0.1 - 0.2 generate moderate shifts pre-test to post-test
- LR = 2 – 5 or 0.5 – 0.2 generate small, but sometimes important changes in probability
- LR = 1 – 2 or 0.5 – 1 are rarely important shifts

MedEd Nicklaus Children's Hospital

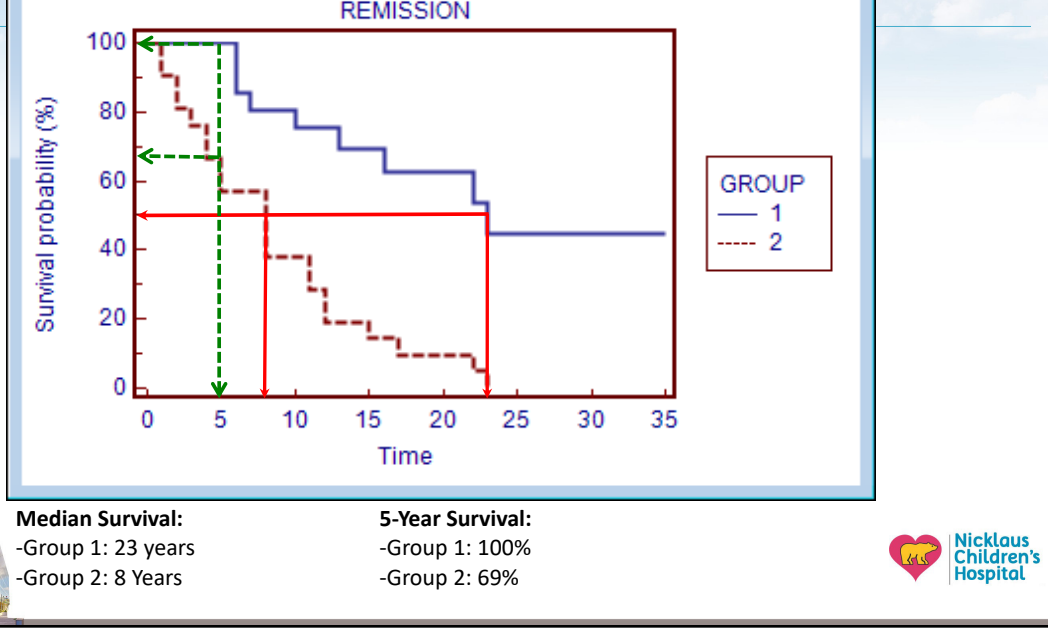
Nicklaus Children's Hospital

64

Prognosis Studies

- RCTs are unethical and not feasible
- Cohort or Case-Control
- How likely are the outcomes over time?
  - Three ways of reporting it:
    - % Survival at a particular point in time (1 year or 5 year survival)
    - Median Survival (Length of F/U by which 50% of the study patients have died)
    - Survival Curves/ Kaplan-Meier Curve (% of study population at each point in time that is free of the specified outcome)

65



Median Survival:
-Group 1: 23 years
-Group 2: 8 Years

5-Year Survival:
-Group 1: 100%
-Group 2: 69%

66

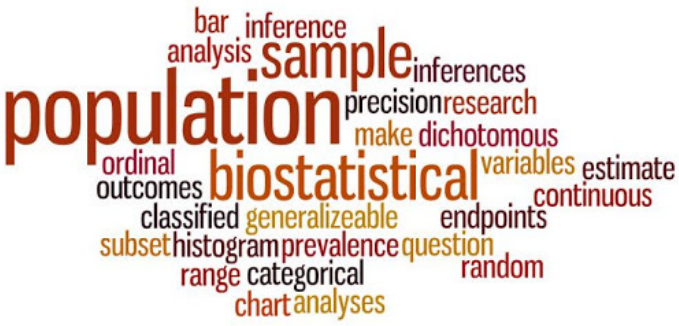
Type of Question and Study Design

Type of Question	Suggested best type of Study
Therapy	RCT > cohort > case control > case series
Diagnosis	Prospective, blind comparison to a gold standard
Etiology/Harm	RCT > cohort > case control > case series
Prognosis	Cohort study > case control > case series
Prevention	RCT > cohort study > case control > case series
Clinical Exam	Prospective, blind comparison to gold standard
Cost	Economic analysis

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

67

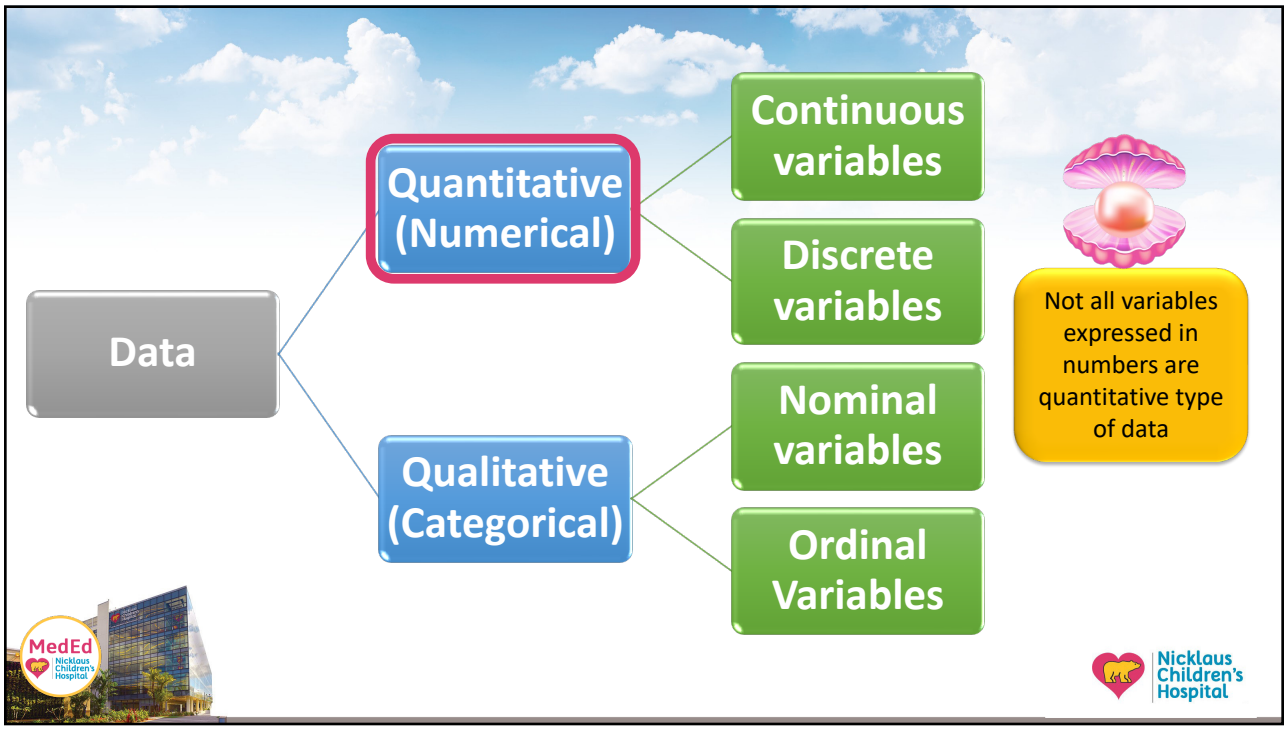


BASIC BIOSTATISTICS

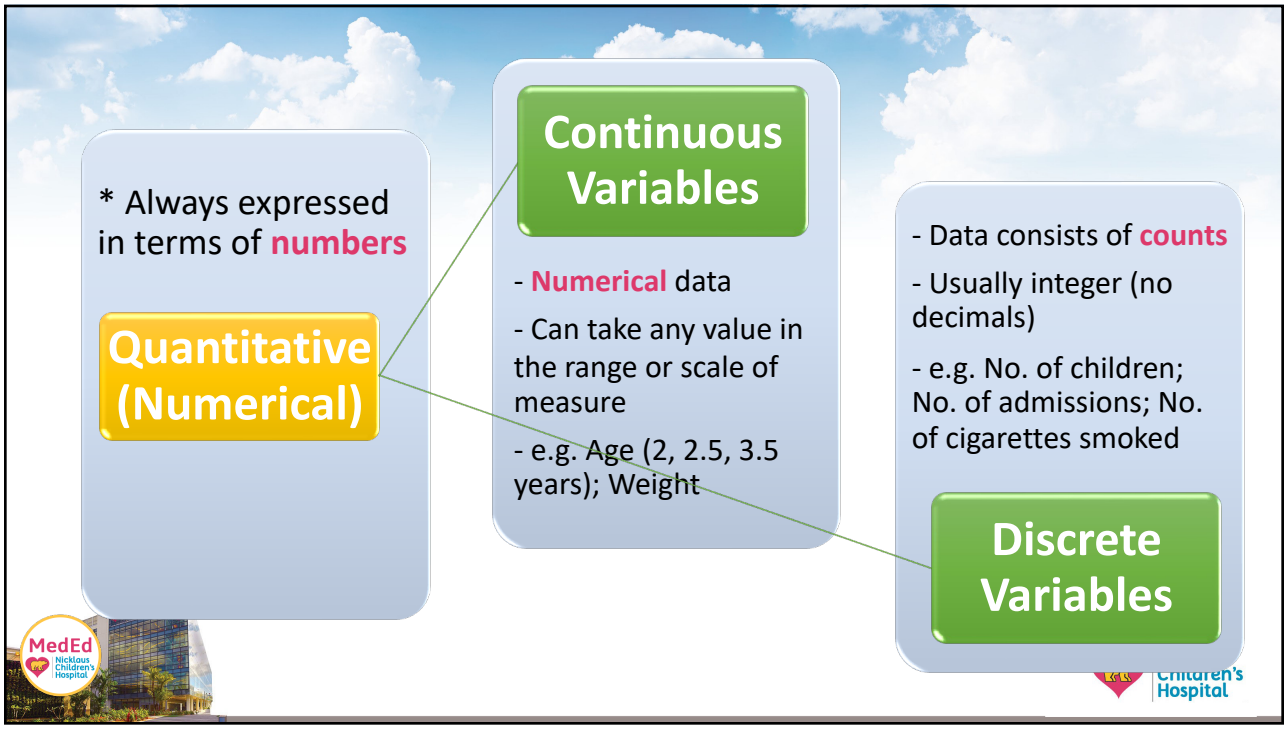
MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

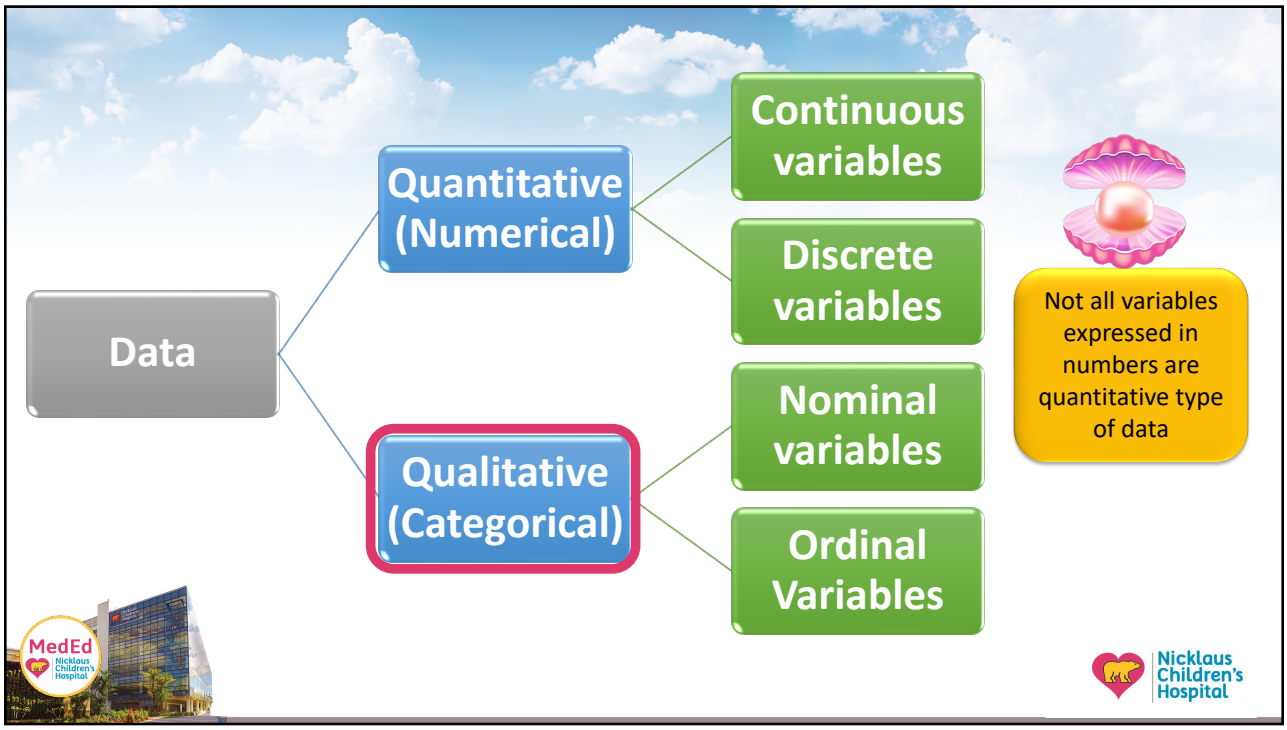
68


Data
Quantitative (Numerical)
Qualitative (Categorical)
Continuous variables
Discrete variables
Nominal variables
Ordinal Variables
Not all variables expressed in numbers are quantitative type of data
MedEd
Nicklaus Children's Hospital

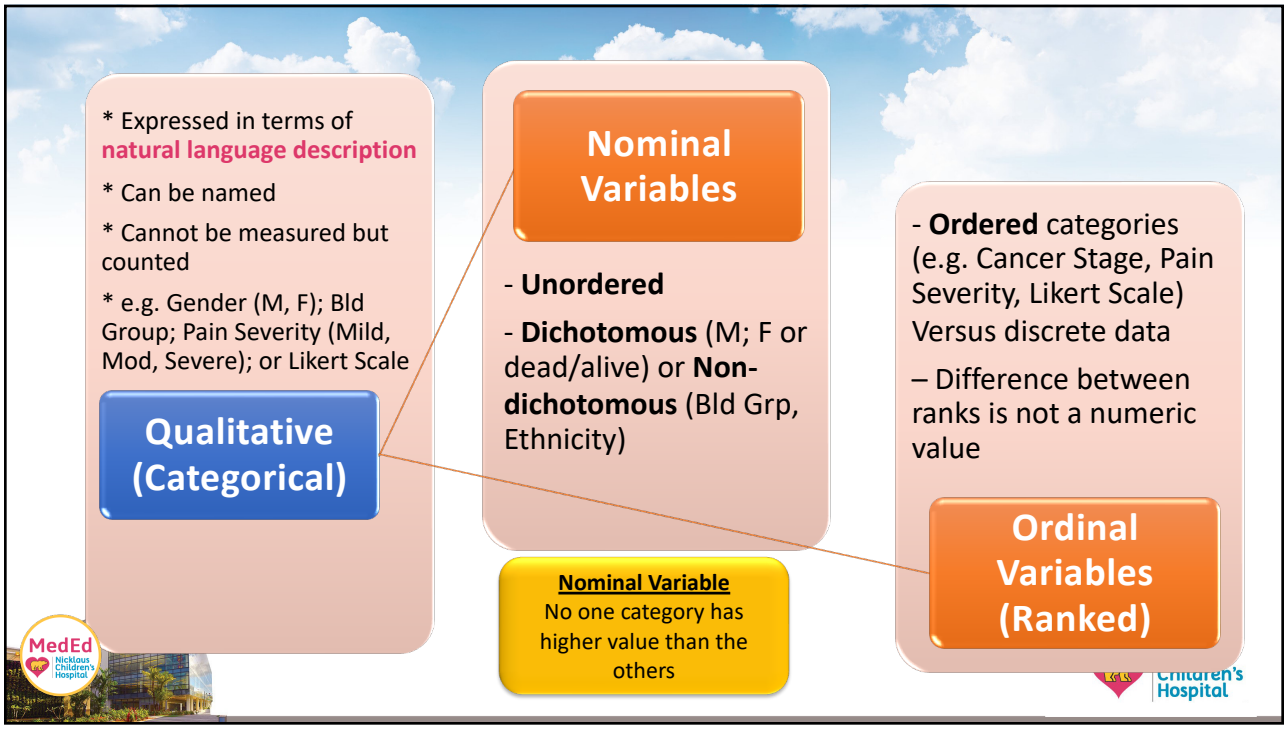
69


* Always expressed in terms of numbers
Quantitative (Numerical)
Continuous Variables
- Numerical data
- Can take any value in the range or scale of measure
- e.g. Age (2, 2.5, 3.5 years); Weight
- Data consists of counts
- Usually integer (no decimals)
- e.g. No. of children; No. of admissions; No. of cigarettes smoked
Discrete Variables
MedEd
Nicklaus Children's Hospital

70


Data
Quantitative (Numerical)
Qualitative (Categorical)
Continuous variables
Discrete variables
Nominal variables
Ordinal Variables
Not all variables expressed in numbers are quantitative type of data
MedEd
Nicklaus Children's Hospital

71


* Expressed in terms of natural language description
* Can be named
* Cannot be measured but counted
* e.g. Gender (M, F); Bld Group; Pain Severity (Mild, Mod, Severe); or Likert Scale
Qualitative (Categorical)
Nominal Variables
- Unordered
- Dichotomous (M; F or dead/alive) or Non-dichotomous (Bld Grp, Ethnicity)
Nominal Variable
No one category has higher value than the others
- Ordered categories (e.g. Cancer Stage, Pain Severity, Likert Scale) Versus discrete data
– Difference between ranks is not a numeric value
Ordinal Variables (Ranked)
MedEd
Nicklaus Children's Hospital

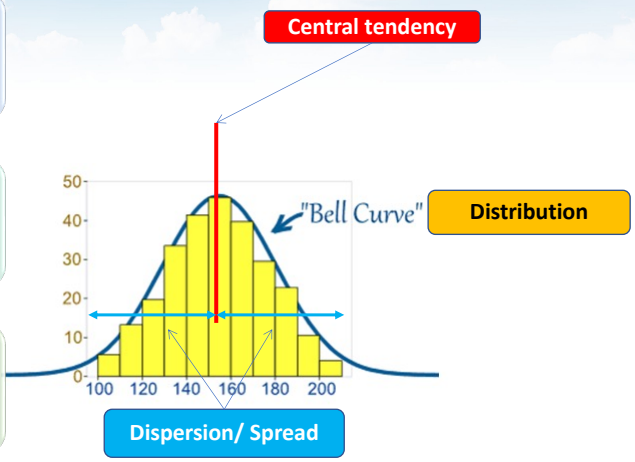
72

Descriptive Statistics – Characteristics include:

Central Tendency
- Estimate of "center" of a distribution of values

Distribution of Data
- Normal (parametric distribution)
- Non-normal (non-parametric)
- Presented as frequency distribution

Dispersion/ Variation
- Spread of values around the central tendency



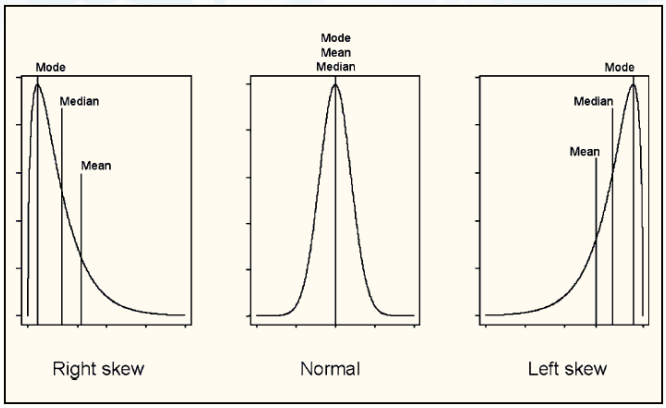
http://www.mathsisfun.com/data/standard-normal-distribution.html

73

Distribution of Data

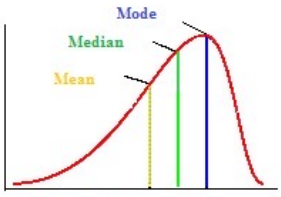
HINT: Mean to the side of the tail to the median

- In a normal distribution:
  - Mean=median=mode
- In skewed to left distribution:
  - Mean<median
    - Eg: birth weight in an NICU
- In skewed to right distribution
  - Mean>median
    - Eg: income

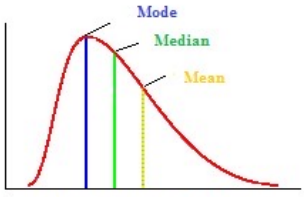


74

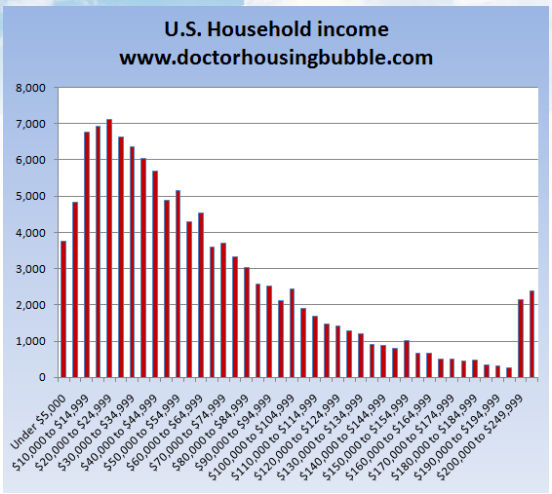
Skewed Distribution



Left-Skewed (Negative Skewness)



Right-Skewed (Positive Skewness)



75

Measures of Central Tendency

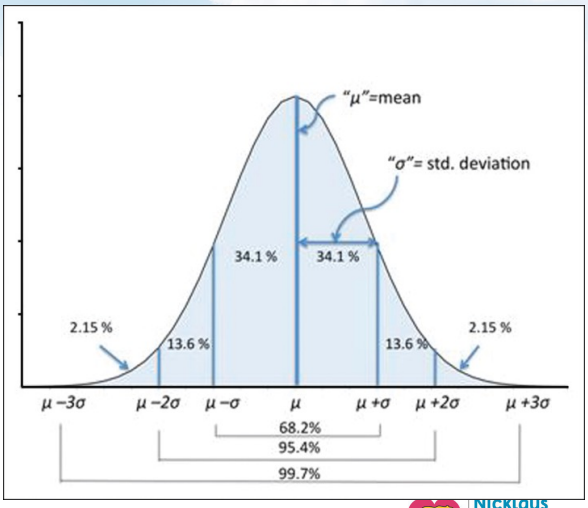
Estimate of center of distribution of values. Three types of estimates

Mean	Median	Mode
• Average of **all** values (uses ALL values in a sample) • **Most commonly used** measure of central tendency • Used in many statistical equations • Influenced by extreme values (Skewed distribution)	• Exact middle of a set of "ordered" values • **Less sensitive to extreme values** • **Better measure of a central tendency in highly skewed distributions** eg. Family income • Should be accompanied by inter-quartile range (IQR) • eg. Family income: median$ 25,000 (25-75 centile range 15,000-45,000)	• most frequently occurring value in a set • It is only measure of central tendency for **nominal data** • Has a high sample fluctuations • A sample may have more than one mode (multimodal distribution) • Eg: O +ve is the most frequent blood group in US

76

Measures of Variation/ Dispersion/ Spread

- Inter-quartile range
  - Around Median
- Variance
- Standard Deviation = the positive square root of the Variance
  - Around Mean
  - 95% sample data with 1.96 SD on each side of the mean
- Range = Largest value – Smallest value



77

Confidence Intervals

- An estimate of a population parameter
- Stated as a range between a lower and upper limit with a specific degree of certainty
- For a given sample size, if you want more confidence that your interval will be correct, you will have a wider interval and therefore, a less precise estimate
- The most commonly used level of certainty is 95%

78

Hypothesis Testing

- A hypothesis = A tentative explanation
  - We seek to prove or disprove the explanation
- Stated as a pair of statements
  - The Null Hypothesis (H_0)
    - A hypothesis which the researcher **tries to disprove, reject, or nullify**
    - True until evidence indicates otherwise
    - If you can conclude that H_0 is false (reject H_0), then the H_1 must be true
  - The Alternative Hypothesis (H_1)
    - Represents the conclusion reached by rejecting H_0
    - We reject H_0 if the evidence from the sample indicates that H_0 is unlikely to be true

Nicklaus Children's Hospital

81

Error Types

Type I Error

- You commit a Type I Error if you **reject a true null hypothesis - H_0**
- Alpha (α) refers to:
  - The risk, or probability, of a type I error occurring
  - Is also known as "level of significance" of the statistical test
  - **You control α** by deciding the risk you are willing to tolerate of a type I error
  - **You specify α before** performing the hypothesis test
  - The most common α values are 0.01 and 0.05

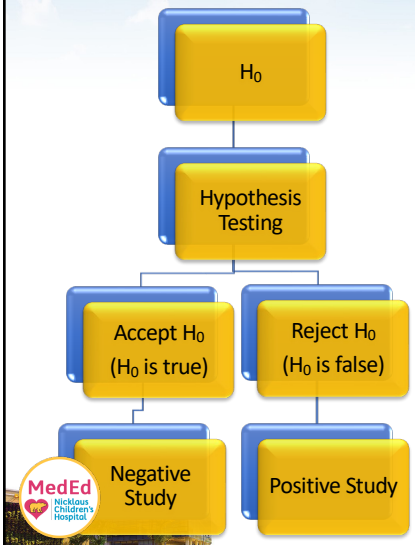
Type II Error

- You commit a Type II Error if you **do not reject a false null hypothesis**
- Beta (β) refers to:
  - The risk, or probability, of a type II error occurring
  - Unlike the type I error, the type II error is **not directly established by you**

Nicklaus Children's Hospital

82

Hypothesis Testing
Probabilities of Type I and Type II Errors



STATISTICAL DECISION (Investigator's Conclusion)	***ACTUAL SITUATION***	
	H_0 IS TRUE	**H_0 IS FALSE**
REJECT H_0	**TYPE I ERROR = α** *(FALSE POSITIVE STUDY)*	**CORRECT DECISION = 1- β (POWER)** *(TRUE POSITIVE STUDY)*
DO NOT REJECT H_0	**CORRECT DECISION = 1- α (CONFIDENCE)** *(TRUE NEGATIVE STUDY)*	**TYPE II ERROR = β** *(FALSE NEGATIVE STUDY)*

Nicklaus Children's Hospital

83

Hypothesis Testing:
Correct & Erroneous Conclusions

STATISTICAL DECISION (Investigator's Conclusion)	***ACTUAL SITUATION***	
	H_0 IS TRUE	**H_0 IS FALSE**
REJECT THE H_0	**TYPE I ERROR = α** *(FALSE POSITIVE STUDY)*	**CORRECT DECISION = 1- β (POWER)** *(TRUE POSITIVE STUDY)*
DO NOT REJECT THE H_0	**CORRECT DECISION = 1- α (CONFIDENCE)** *(TRUE NEGATIVE STUDY)*	**TYPE II ERROR = β** *(FALSE NEGATIVE STUDY)*

Type I Error

- Erroneously concluding H_0 to be false
- Rejecting H_0 when it is true
- False positive study

Type II Error

- Erroneously concluding H_0 to be true
- Accepting H_0 when it is false
- False negative study

α level

- The chance (probability) of rejecting H_0 when it is true is α level
- **The chance of Type I error is α level**
- The chance of false positive study is α level

β level

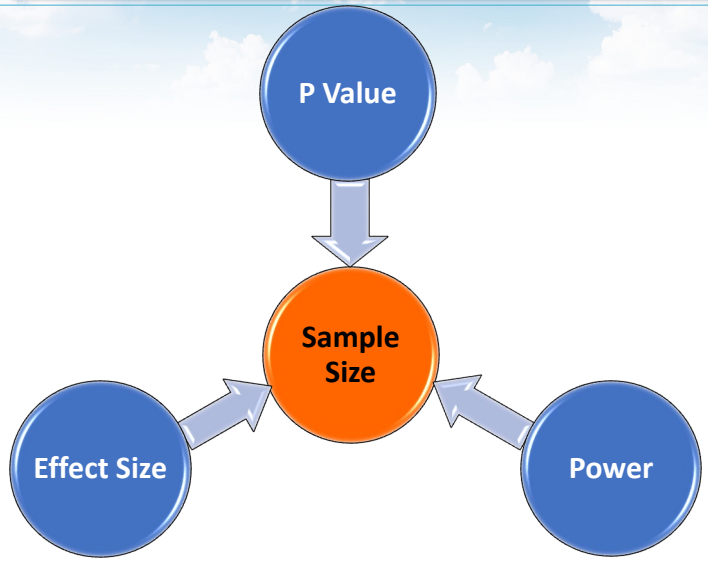
- The chance (probability) of accepting H_0 when it is false is β level
- **The chance of Type II error is β level**
- The chance of false negative study is β level

84

Sample Size

Effect size

The smaller the effect size, the larger the sample size required



85

The p-Value

- The actual risk of having a **type I error**
- AKA the **"observed"** level of significance
- Represents the chance of detecting a difference (inequality) in the parameters by chance when in fact there is no difference at all
- There is not a firm division between what scientists consider true and not true, but traditionally a **p-value of 0.05 or less** has been accepted as evidence of actual difference
- If p were 0.05 this means there is one chance in 20 that you could detect a difference (rejected H_0) by pure chance when in reality there was no real difference (H_0 is true)

**** α is set by investigators ****

**** p-value is calculated ****

86

Hypothesis Testing

- Significance tests are carried out on the **assumption that H_0 is true**
- An α level is set = **probability of rejecting a null hypothesis when it is true.**
  - **False positive study**
- This α level is same as p-value of significance.
  - Usually the α level is set at 0.05 (5%).
- When H_0 is rejected with a p-value of <0.05, we can conclude that the rejection is unlikely due to chance alone (<5%).

The smaller the set p-value (α level); the larger the sample size required

87

Power

- Power = Probability of correctly identifying the difference between the two groups in study sample when one genuinely exists in population from which the samples are drawn.
  - **Probability of (true) positive study**
- Power is 1-β
  - β is the probability of Type II error
- The sample size depends upon the power of the study.

The higher the power (1-β), the larger the sample size required, and the lower the probability of Type II error

88

Common Statistical Tests

Data	Numerical (parametric)	Numerical (non-parametric) Ranks, Scores	Binomial (2 X 2)
Describe one group	**Mean with Standard deviation**	**Median with Inter quartile range**	**Proportion or %**
Compare two unpaired groups	**Unpaired t-test**	**Mann-Whitney Test**	**Chi-square (Fisher's ≤5)**
Compare two paired groups	**Paired t-test**	**Wilcoxon test**	McNemar's test
Compare ≥3 unmatched groups	**One-way ANOVA**	**Kruskal-Wallis test**	**Chi-square**
Compare ≥3 matched groups	Repeated-measures **ANOVA**	Friedman test	
Association between 2 variables	**Pearson correlation**	**Spearman correlation**	
Predict value from another variable	Simple linear (non-linear) **regression**	Non-parametric **regression**	Simple logistic **regression**
Predict value from several variable	Multiple linear (non-linear) **regression**		Multiple logistic **regression**

89

Chi-square *vs* Fisher's

- Fisher's exact test is more accurate
  - Difficult to calculate manually for a large data
  - If the data points are <6 (≤5) in any cell (of 2X2 Table) always use Fisher's
- Chi-square test
  - If the data points are >5 (results will be same as Fisher's test)
  - For comparing 3 or more groups use Chi-square

90

T-test *vs* ANOVA

- T-test is used to compare the means of TWO groups
- Paired t-test is used when the data is matched
  - eg: pre and post data
- ANOVA is to compare the means of THREE or more groups
  - ANOVA = analysis of variance
- Both are used for parametric data only
  - Parametric = normally distributed

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

91

Phases of Clinical Trials

MedEd Nicklaus Children's Hospital

Nicklaus Children's Hospital

92

Phases of Clinical Trials

Phase	1ary Goal	Dose	Patient Monitor	Typical No. of participants	Notes
Preclinical	• Testing in non-human subjects • Gather efficacy, toxicity and pharmacokinetic info	Unrestricted	Graduate level researcher (PhD)	N/A (In vitro and in vivo only)	
Filing & Approval of IND* (Investigational New Drug) Application					
Phase 0	• Pharmacodynamics and Pharmacokinetics • Particularly oral bioavailability and half-life of the drug	Very small, subtherapeutic	Clinical researcher	10 people	Often skipped for phase I
Phase I	Testing of drug on healthy volunteers for dose ranging	Often subtherapeutic, but with ascending doses	Clinical researcher	20-100	• Determine effectiveness • Evaluate Safety
Phase II	Testing of drug on patients to assess efficacy and safety	therapeutic dose	Clinical researcher	100-300	• Determines efficacy • At this point, the drug is not presumed to have any therapeutic effect whatsoever

*IND – Permission to conduct the Clinical Trial

93

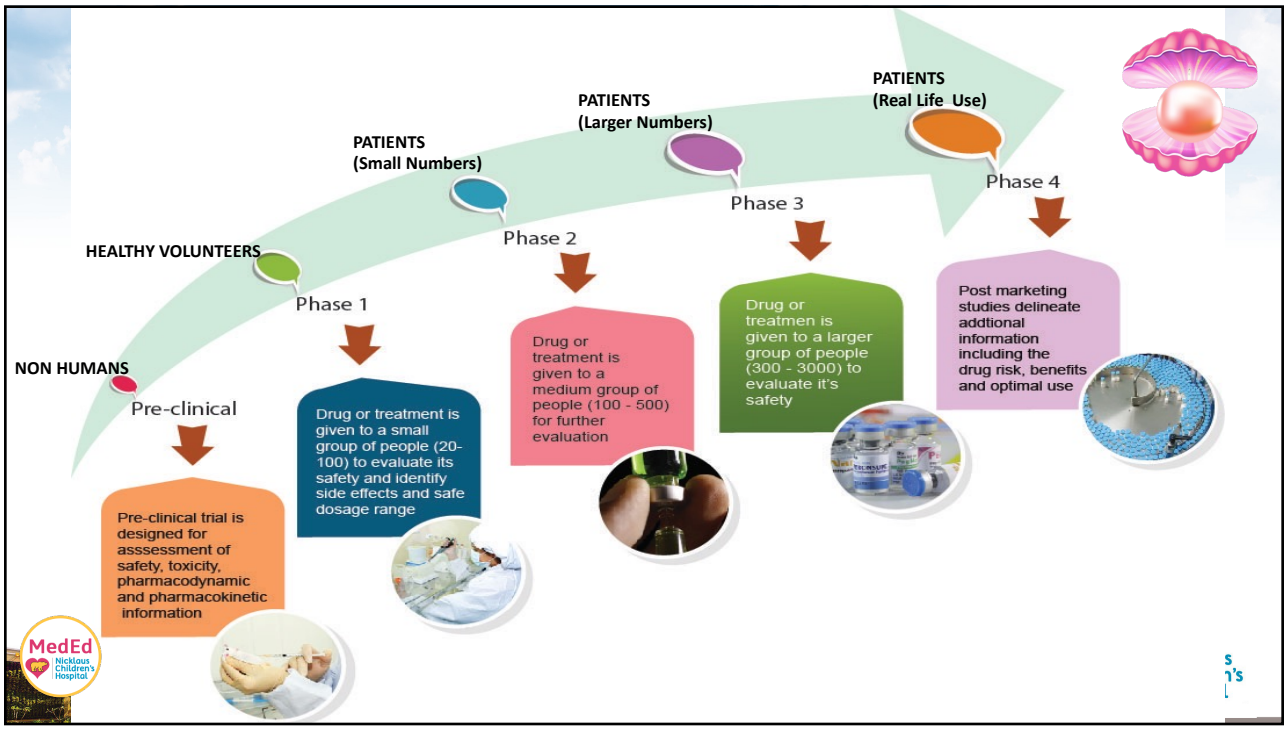
Phases of Clinical Trials

Phase	1ary Goal	Dose	Patient Monitor	Typical No. of participants	Notes
Phase III	• Testing of drug on patients • Assess efficacy, effectiveness and safety	Therapeutic dose	clinical researcher and personal physician	1000-2000	• Determines a drug's therapeutic effect • At this point, the drug is presumed to have some effect • Confirm effectiveness • Monitor side effects • Compare it to standard treatment • Collect info to use the drug safely
Filing and Approval of NDA* (New Drug Application) to FDA to Approve the Drug for Marketing					
Phase IV	Postmarketing Surveillance – watching drug use in the public	Therapeutic dose	personal physician	Anyone seeking treatment from their physician	• Watch long-term effects & side effects • Info on drug effect in various populations
Phase V	Translational research	No dosing	None	All reported use	Research on data collected

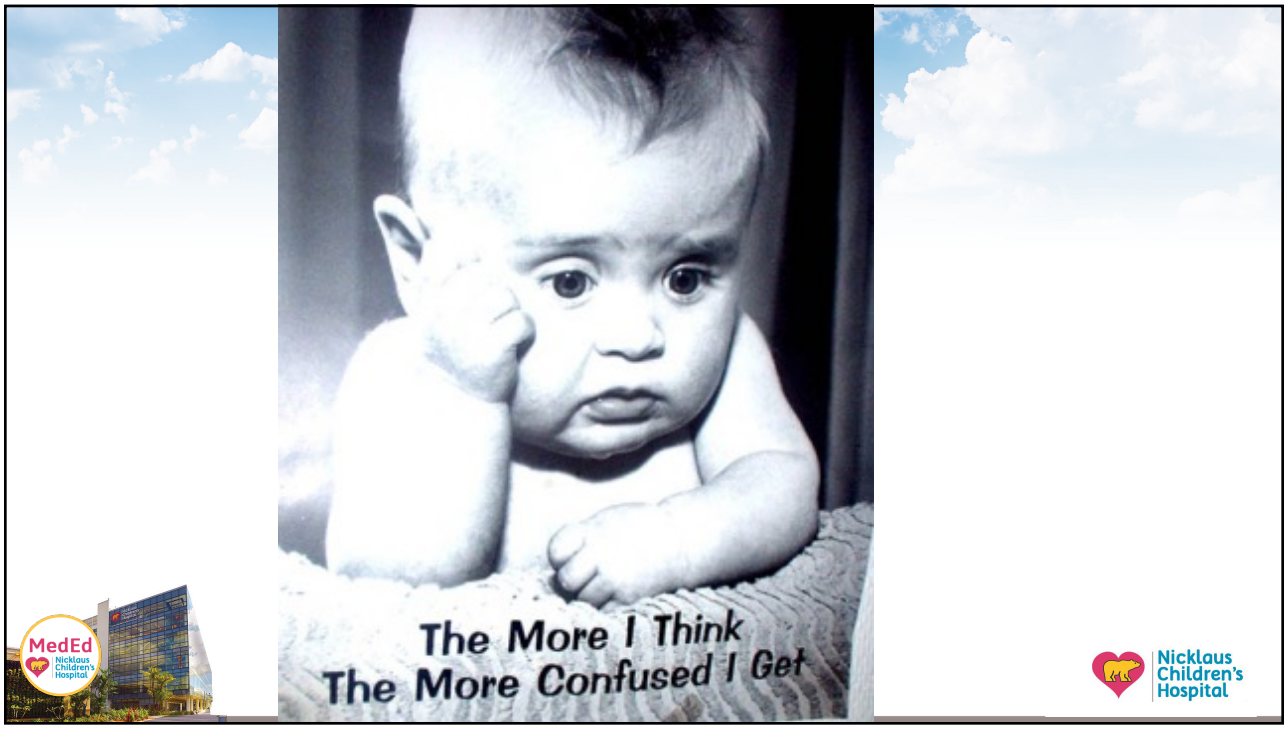
lm.nih.gov/services/ctphases.html
pedia.org/wiki/Phases_of_clinical_research

*NDA – Permission to Market the New Drug

94


PATIENTS
(Real Life Use)
PATIENTS
(Larger Numbers)
PATIENTS
(Small Numbers)
HEALTHY VOLUNTEERS
NON HUMANS
Phase 4
Phase 3
Phase 2
Phase 1
Pre-clinical
Post marketing studies delineate addtional information including the drug risk, benefits and optimal use
Drug or treatmen is given to a larger group of people (300 - 3000) to evaluate it's safety
Drug or treatment is given to a medium group of people (100 - 500) for further evaluation
Drug or treatment is given to a small group of people (20-100) to evaluate its safety and identify side effects and safe dosage range
Pre-clinical trial is designed for asssessment of safety, toxicity, pharmacodynamic and pharmacokinetic information
MedEd
Nicklaus Children's Hospital

95


The More I Think
The More Confused I Get
MedEd
Nicklaus Children's Hospital
Nicklaus Children's Hospital

96


Nicklaus Children's Hospital
MedEd
Nicklaus Children's Hospital
Спасибо
Gracias
شكرا
Obrigado
Спасибо
Dank U
Grazie
Ευχαριστώ
Danke
Merci
Obrigado
Diolch
Thank You
Ngiyabonga
Dank U
Thank
You
Merci
Ngiyabonga
Tack
Dzięki
Danke
Grazie
תודה
Dank U
Diolch
Terima Kasih
Grazie
Tack
Ευχαριστώ
Dzięki
Diolch
Gracias
Merci
תודה
Tack

97