2-18-2021

Epidemiology, Biostatistics, and Population Health Concepts

Allison R. Casola, PhD, MPH, MCHES
Thomas Jefferson University

SIDNEY KIMMEL MEDICAL COLLEGE

Epidemiology, Biostatistics, and Population Health Concepts

Allison R. Casola, PhD, MPH, MCHES
Thursday, February 18, 2020
Resident Conferences, Dept. Family and Community Medicine
Thomas Jefferson University

Jefferson
Thomas Jefferson University

Overview

Study Designs

- Overview

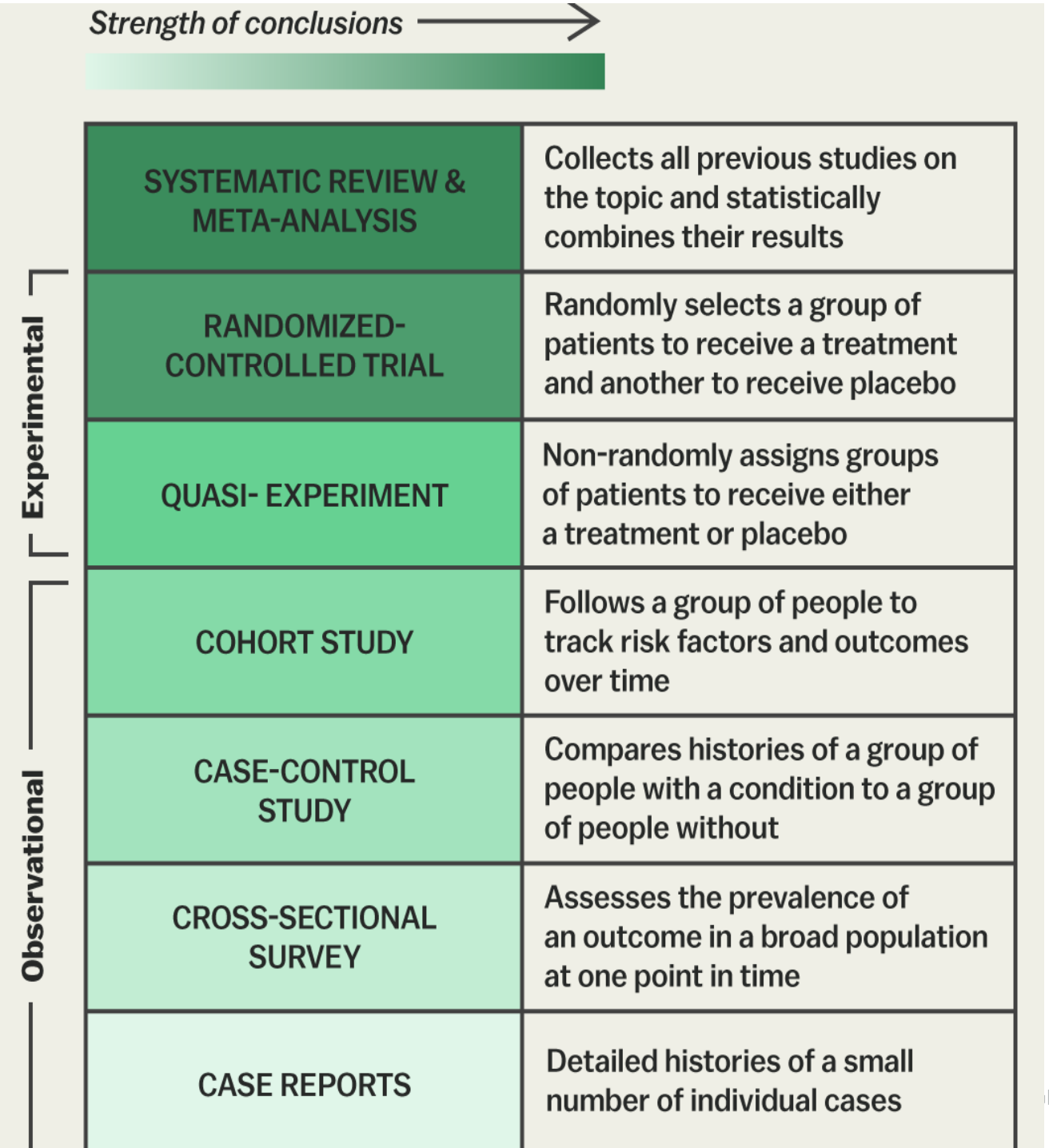
Epidemiology and Biostatistics

- Vocabulary Review
- Reliability, Variability, and Bias
- Diagnostic Testing and Measures of Association

Population Health

- Evidence-Based Medicine
- Interpreting and Evaluating Literature
- Clinical Decision-Making

Study Designs

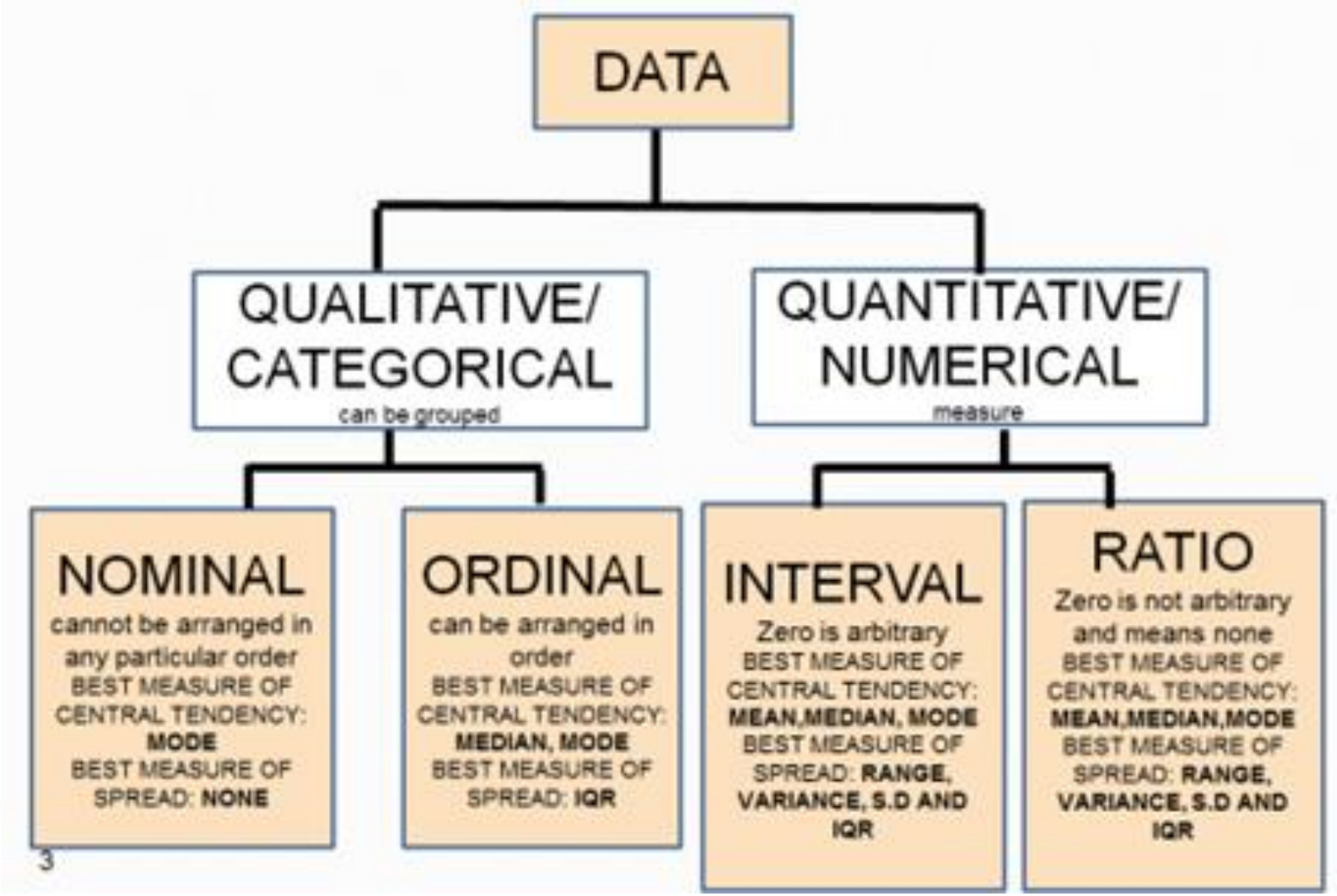


What is a study design?

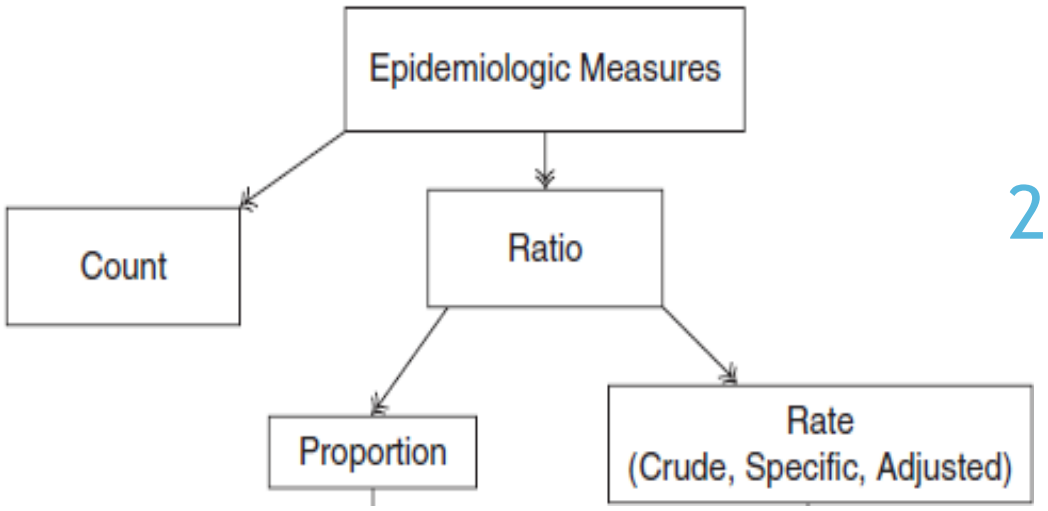
- Framework, or set of methods and procedures, used to collect and analyze data on specific variables denoted by a research problem
- Types of design employed to answer different types of research questions
- Design types have strengths and limitations

Epidemiology and Biostatistics

Variable Scales Measurement



Counts, Ratios, Rates, and Proportions



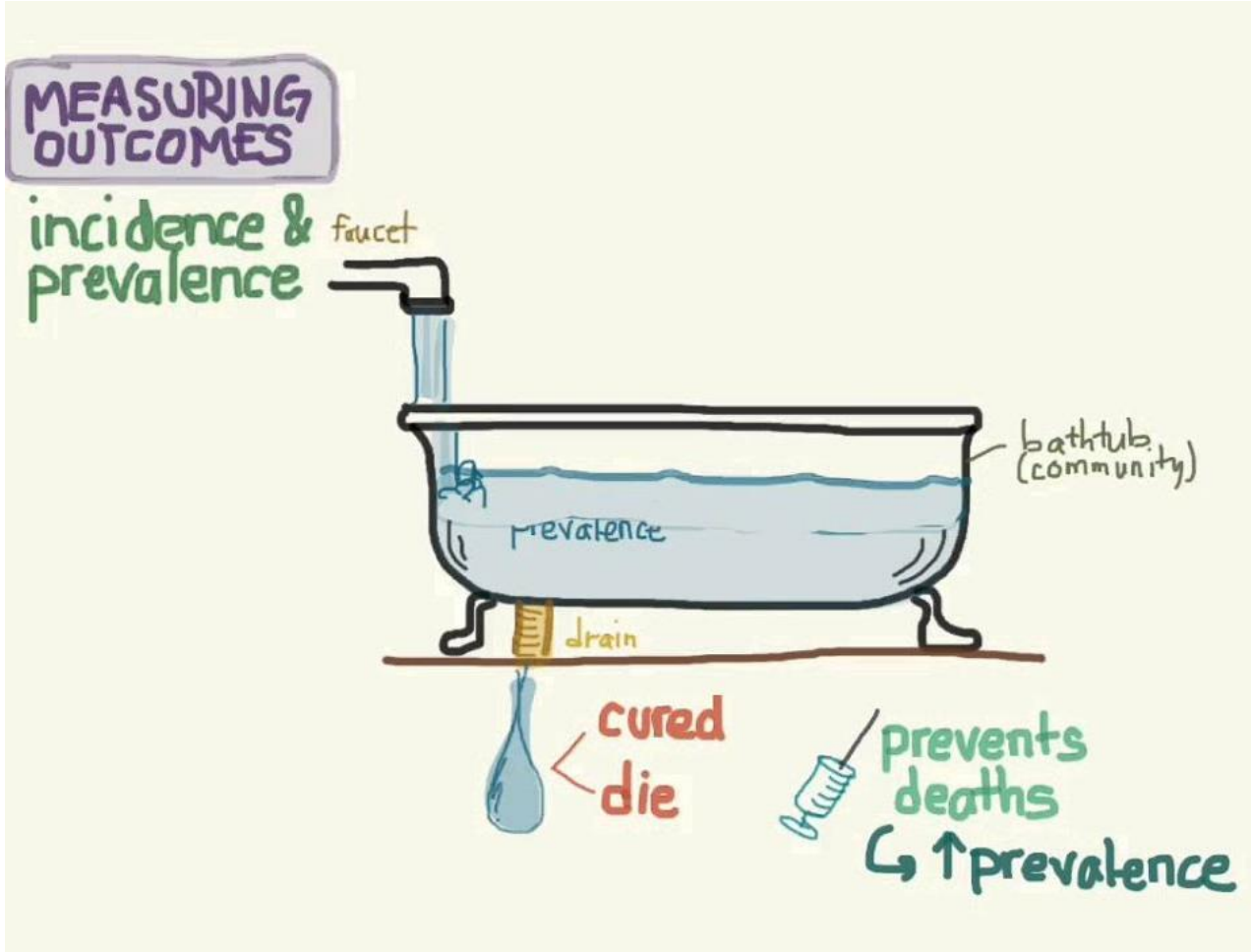
1. Count
   - Simplest and most frequently performed quantitative measure
   - Number of cases of a disease or health phenomenon

2. Ratio
   - Division of one quantity by another (fraction!)

   A. Proportions
      - Ratio where numerator is part of denominator.
      - May be expressed as a *percentage (Proportion * 100%)*

   B. Rates (*risk!)*
      - Ratio where time forms part of the denominator.
        - Disease frequency, time period, population unit size

Incidence and Prevalence



- Incidence = new cases
  - Rate of (# new cases) / (# people at risk in given time frame)
  - Measuring how fast the disease is occurring = risk
- Prevalence = all cases (in total)
  - Proportion (%) of (# cases) / (# total people)
  - How much disease is in the population? = burden of disease

Validity and Reliability

- Validity = degree measurement reaches correct conclusion
  - (1) Internal Validity
    - Results accurately reflect true situation of study population, defined by boundaries of the study
    - Provides a true estimate of effect, given the limits of the population studied
  - (2) External Validity
    - Results are applicable to other populations (generalizability)
    - Do these results apply to other patients, such as patients who are older, sicker, or less economically advantaged than subjects in the study?

- Reliability = degree to which an instrument can produce precise, consistent results, repeatedly over different occasions, with there is no evidence of change.

Bias

- Can be random or systematic
  - Systematic error that leads to distortion of results
- Can occur anywhere; big concern in observational studies because of lack of randomization
- (3) categories of systematic bias
  1. Selection bias
  2. Information bias
  3. Confounding

1. Selection bias
   - Distortion because of how subjects are sampled/selected
   - The selection process increases or decreases chance a relationship between exposure and outcome will be detected
     - Volunteers may differ from non-volunteers
     - Selecting participants based on referral to clinical facilities
       - Potential disconnect with diagnosis and subsequent eligibility
     - Hospitalized v. non-hospitalized patients

Bias (cont.)

2. Information Bias (aka misclassification)

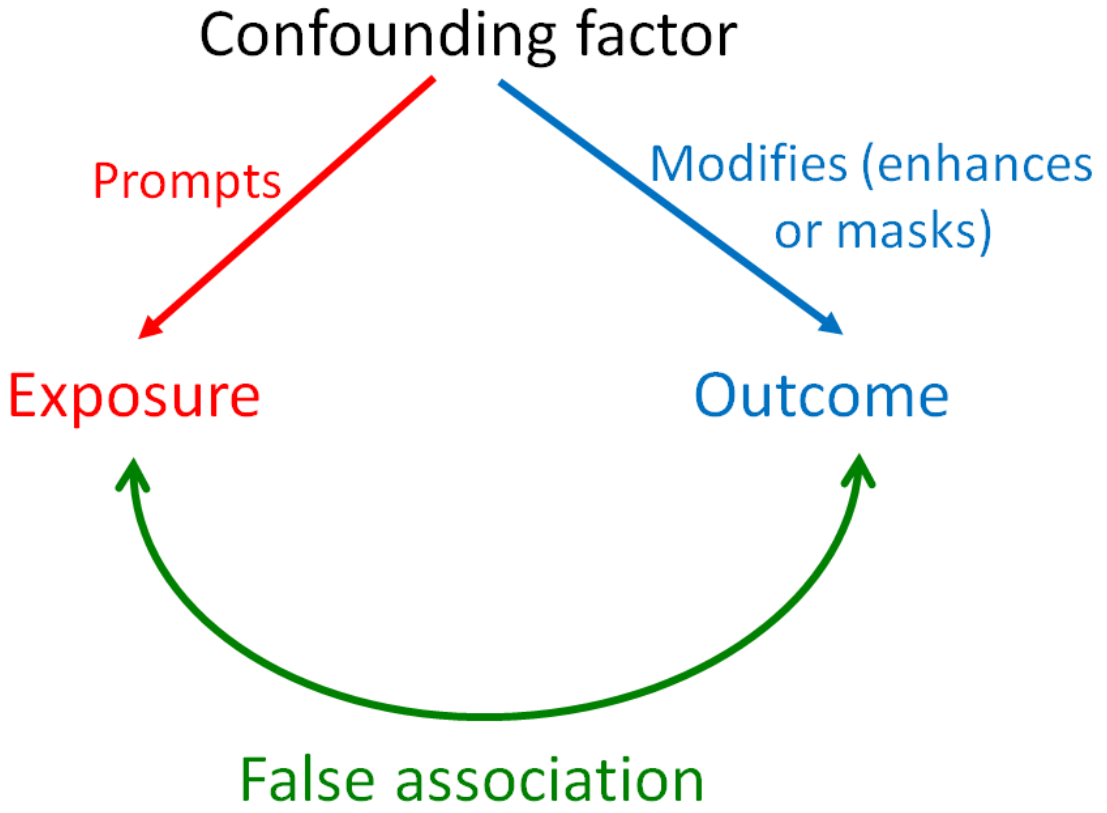
- Distortion because of manner data are collected or classified
  - Random or systematic inaccuracy of measurement
    - Investigator may assign correctly or incorrectly, leading to increase in true positives, true negatives, false positives, false negatives
- (2) Types:
  - *Nondifferential:* errors in classification of exposure/disease that do not depend on the other variable; underestimates the true association
  - *Differential:* errors in classification of exposure/disease that depends on the other variable; may over or underestimates the true association
    - Recall bias: differential ability of subjects to remember previous activities and exposures
    - Interviewer bias: results may be influenced by how the interviewers collect information

Bias (cont.)

3. Confounding

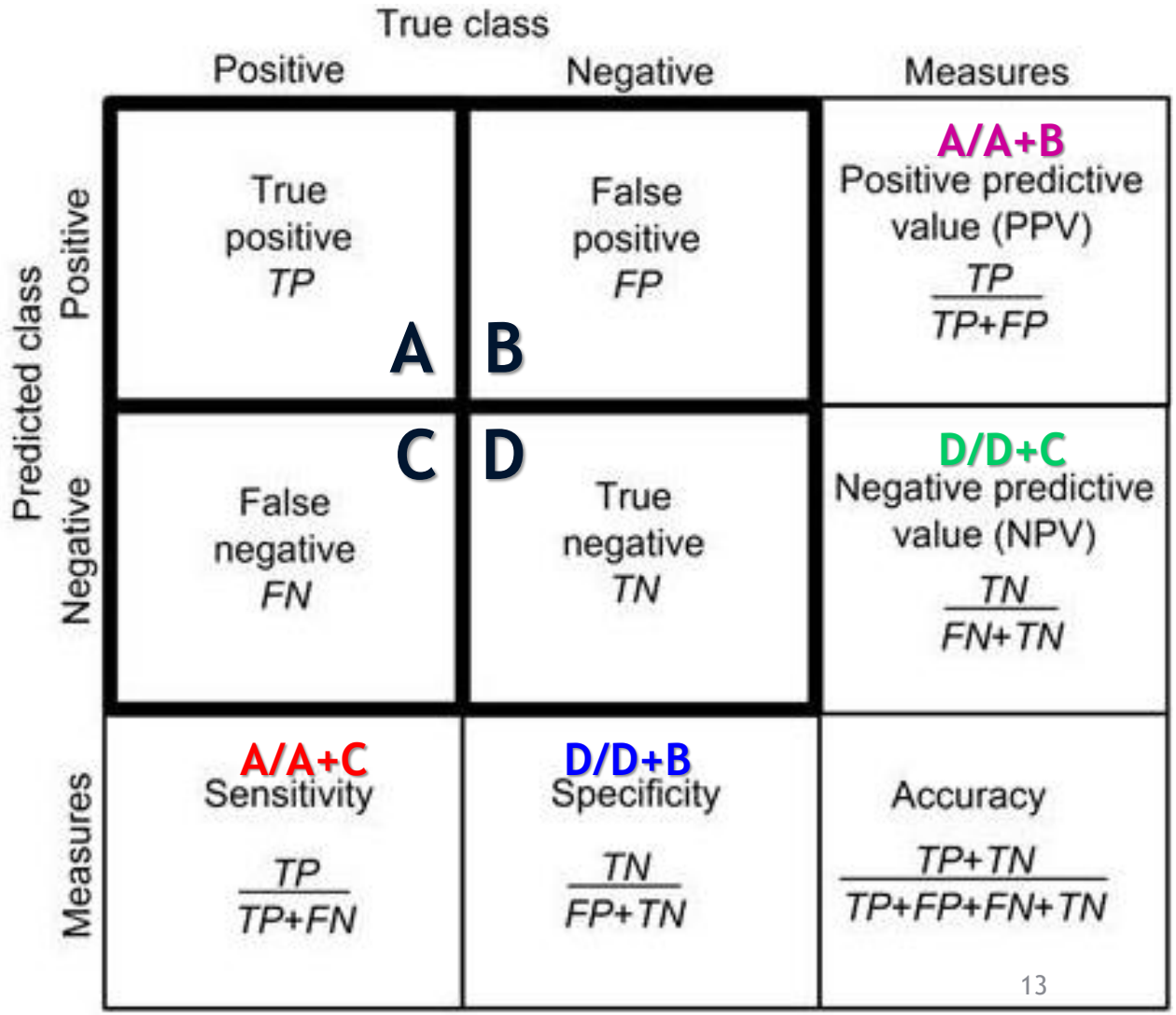
- Extraneous factors distorts true association being examined
- Mixing of the effect of an extraneous "lurking" variable with the effects of exposure and disease

- Confounding Requirements:
  1. The confounding factor must be associated with the outcome
  2. The confounding factor must be associated with the exposure
  3. A confounder cannot be an intermediary step in the causal pathway from the exposure to the outcome



Determining Validity of a Clinical Test

- Descriptors of Test Accuracy
  - Sensitivity (SNout)
    - Proportion of persons with disease who test positive (a/a+c)
  - Specificity (spIN)
    - Proportion of healthy persons who test negative (d/d+b)
- Probability Estimates:
  - Positive Predictive Value (PPV)
    - Proportion of persons with positive test who actually have disease (a/a+b)
  - Negative Predictive Value (NPV)
    - Proportion of persons with negative test who do not have disease (d/d+c)



Measures of Association: Odds Ratio (OR)

- Odds that outcome will occur given exposure v. odds of outcome occurring in absence of exposure
- ratio of two odds
  - outcome $_{exposed}$/outcome$_{unexposed}$
- What OR means…
  - OR = 1.0: no difference
  - OR > 1.0: increase in odds
  - OR < 1.0: decrease in odds

Odds ratio calculation

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

Exposure	Cancer ✔	Cancer ✖
✔	a	b
✖	c	d

Example

$$OR = \frac{354 / 143}{293 / 511}$$ where

Exposure	Cancer ✔	Cancer ✖
✔	354	143
✖	293	511

$$OR = 4.32$$

The odds of lung cancer is 4.32 times greater among smokers compared to non-smokers.

Measures of Association: Relative Risk (RR)

Relative risk calculation

$$RR = \frac{a / (a + b)}{c / (c + d)}$$ where

Exposure	Cancer ✔	Cancer ✖
✔	a	b
✖	c	d

Example

$$RR = \frac{354 / (354 + 143)}{293 / (293 + 511)}$$ where

Exposure	Cancer ✔	Cancer ✖
✔	354	143
✖	293	511

$$RR = 2.54$$

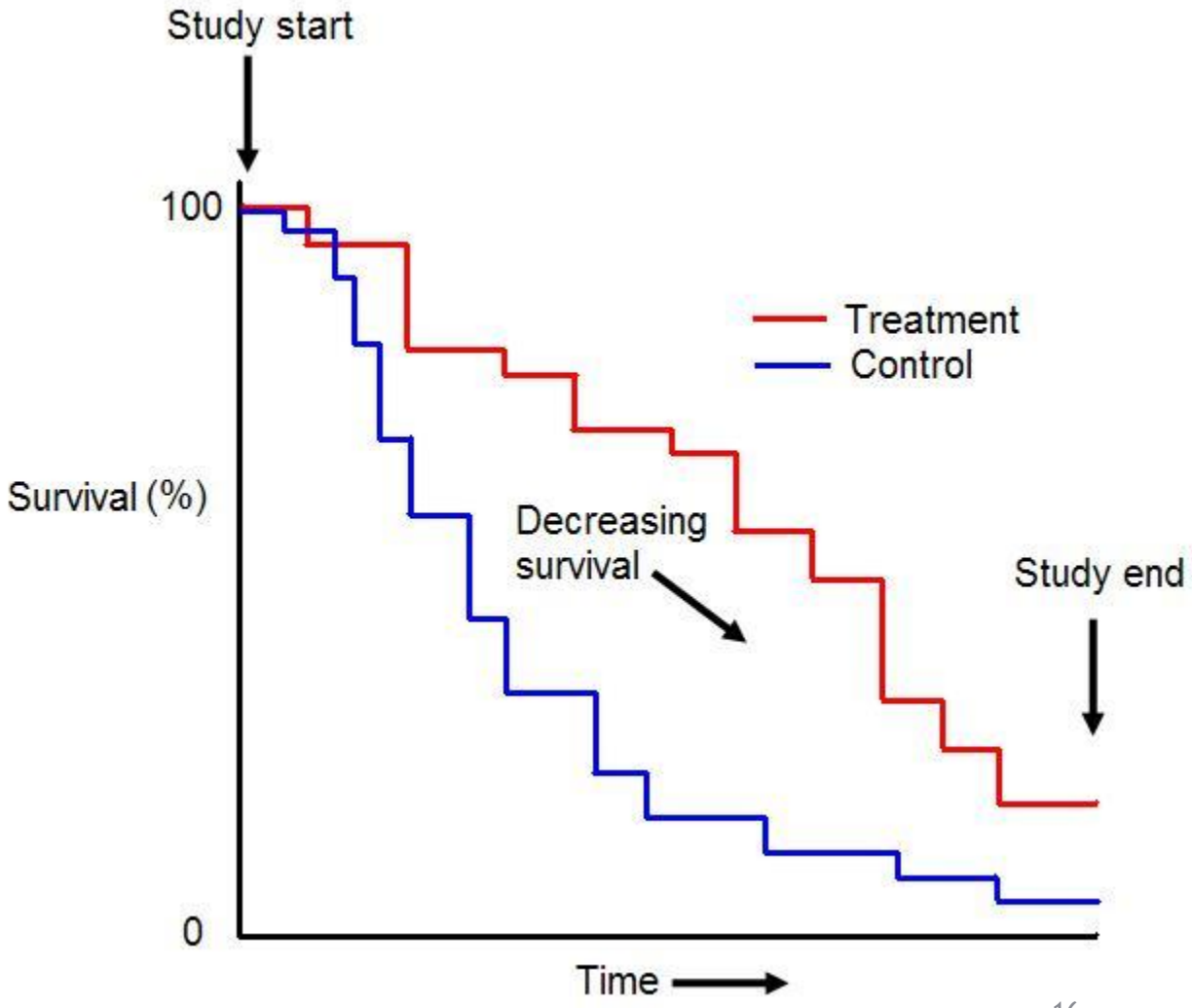
The risk of developing lung cancer is 2.54 times greater among smokers compared to non-smokers.

- Likelihood that an exposed group will develop an outcome relative to those unexposed
  - Compares the risk of a health event among one group with the risk among another group
- AKA “risk ratio” or “rate ratio”
  - $risk_{exposed} / risk_{unexposed}$

Measures of Association: Hazard Ratio (HR)

- Measure of effect of intervention on outcome over time
  - Reported as time-to-event analysis or survival analysis
    - Type of relative risk used to express treatment effects, determine how groups change relative to each other (faster/slower)
  - Rate in which events happen in treatment v. control group as a function of time
    - $hazard_{intervention}$ / $hazard_{control}$

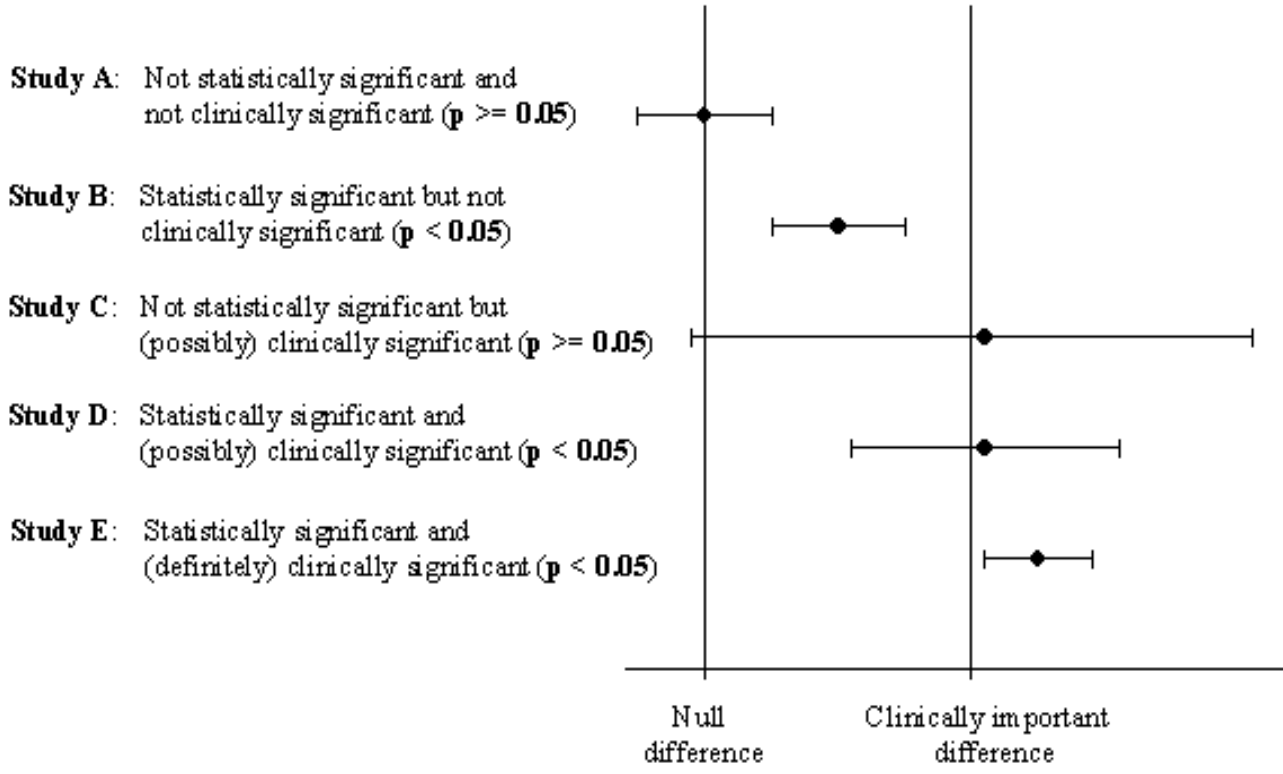
- Kaplan-Meier curves
  - Relative event rate in the groups
  - Y-Axis = prob. of survival (%); X-Axis = time
  - Event hazard = slope of the graph (events/time)



P-Values

P-Value

- Probability that observed result is due to chance alone
- "Is there a statistically significant difference between treatments/groups?
  - Interpretation based on cutoff/ level of significance (e.g. 0.05, 0.01)
- Observed association gives no indication about clinical importance
  - Information from single sample will always leave some level of uncertainty
  - *Confidence intervals!*



Confidence Intervals (CI)

- Range of values for a measure that is believed to contain the true value; within specified certainty level
  - Estimation, magnitude and direction of difference
  - Range of values within which we are reasonably confident that the true association (OR, RR, HR = point estimate)
    - E.g. OR = 2.5; 95%CI: 1.7 - 3.6

- Example: Imagine the study is repeated thousand times
  - About 95% of time the different possible results obtained will lie in this interval.
    - Therefore, we say we are 95% confident that the true population value of what we are estimating in our study lies within the interval.

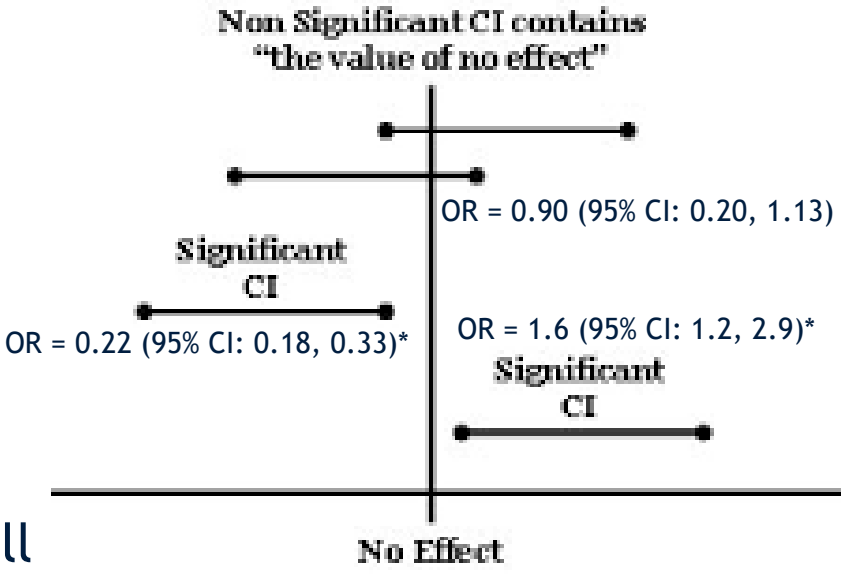
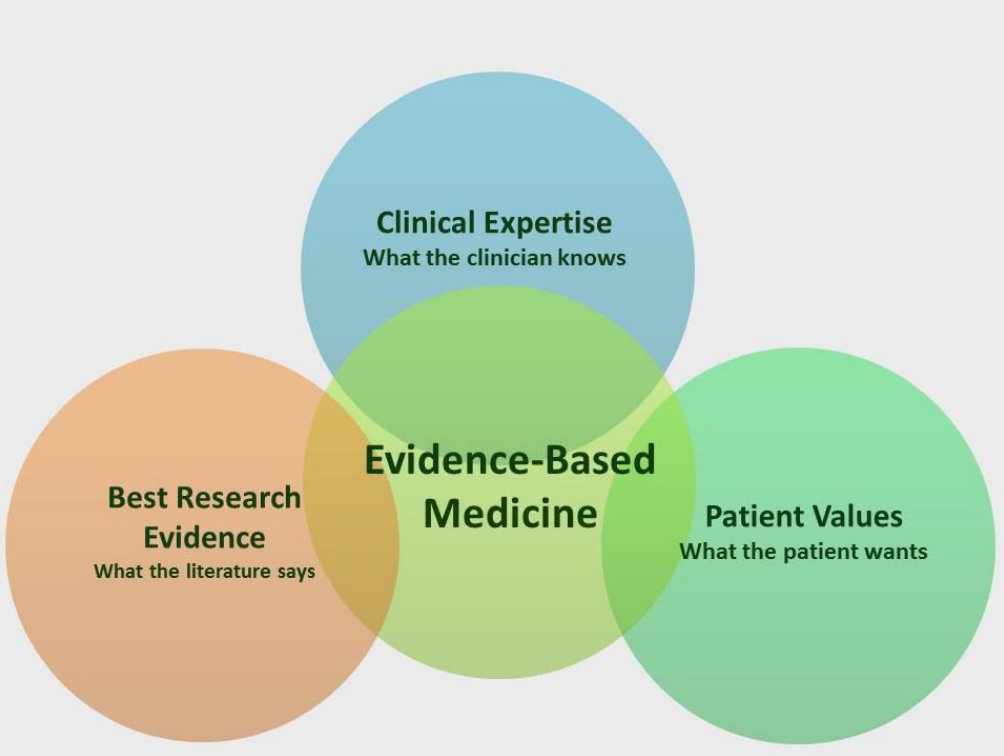


Fig. 1: significant & non significant CIs

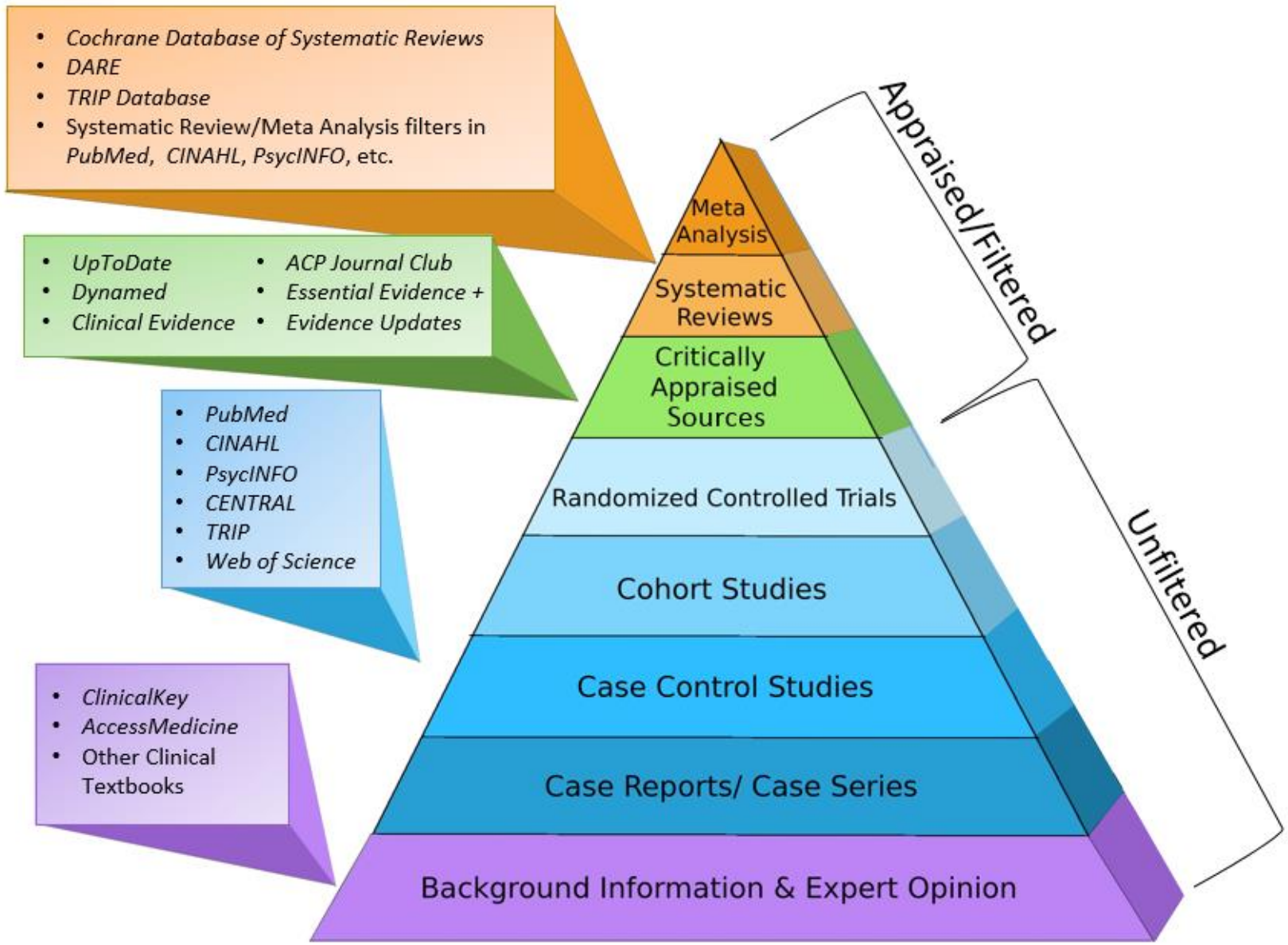
Population Health

Evidence-Based Medicine



- Making health care decisions based on:
  - Current best evidence
  - Clinical expertise
  - Pathophysiological knowledge
- Provide objective basis for selecting strategy for disease management
  - Serve as basis for modifying practice
    - Facilitates assessment
    - Identifies gaps in knowledge
    - Suggest opportunities for improved care quality
- *But what is "evidence"?*

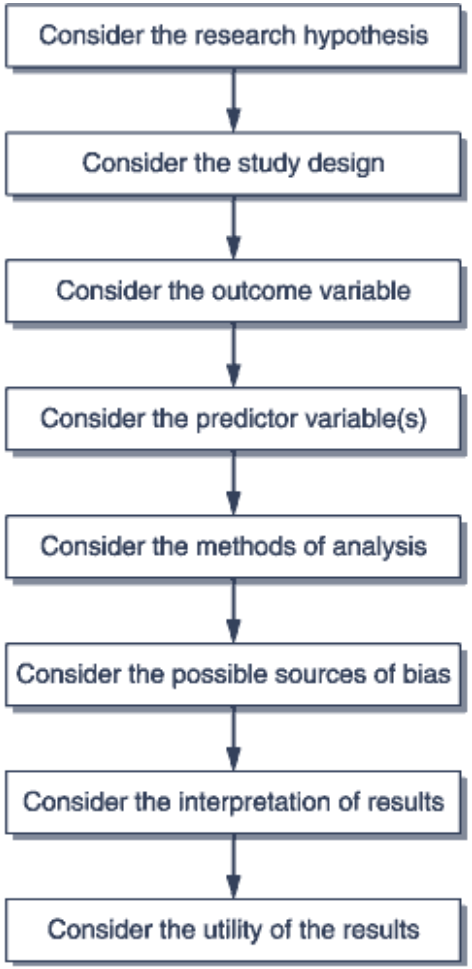
Levels of Evidence



Assessing and Evaluating Literature

- Evaluation of an individual study should include:
  - Research hypothesis
  - Study design
  - Variables
  - Methods of analysis
  - Sources of bias

- Useful to apply a uniform and thorough approach to evaluating the articles



Source: Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR: *Medical Epidemiology*, 4th Edition: http://www.accessmedicine.com

STROBE Statement

Strengthening the reporting of observational studies in epidemiology

u^b UNIVERSITÄT BERN

- Home
- Aims
- News
- Available checklists
  - Previous checklists
- Publications
- Translations
- Commentaries
- Discussion forum
- STROBE group
- Endorsement

STROBE checklists

Version 4 as published in Oct / Nov 2007!

- STROBE checklist for **cohort, case-control, and cross-sectional studies** (comb
  download PDF / Word
- STROBE checklist for **cohort, case-control, and cross-sectional studies**
  download PDF / Word
- Checklist for **cohort studies**
  download PDF / Word
- Checklist for **case-control studies**
  download PDF / Word
- Checklist for **cross-sectionalstudies**
  download PDF / Word
- Draft STROBE checklist for **conference abstracts**
  download PDF

STROBE: *Strengthening the Reporting of Observational Studies in Epidemiology*

- 22-item checklist authors needs to fulfil before submitting manuscript
- Guidelines created to aid authors in ensuring high-quality presentation of observational studies

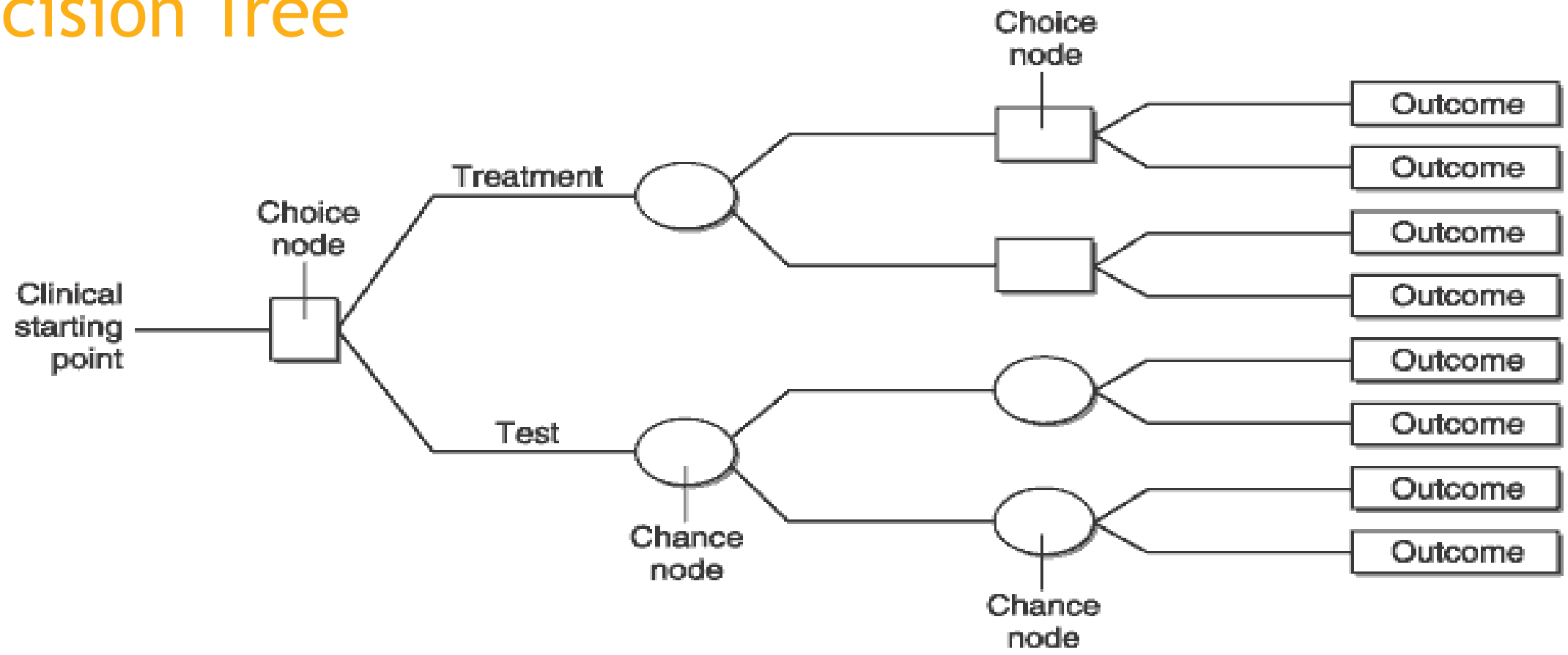
STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable

Clinical Decision-Making

- Understanding of the derivation, calculation, and use of probabilities is important in making clinical decisions
- Formal decision analysis = explicit process, uses information from epidemiologic studies to determine course of action
  - Elements of formal decision analysis:
    - 1. Decision tree diagram
    - 2. Probabilities for uncertain events
    - 3. Test results
    - 4. Alternative outcomes

Decision Tree



Source: Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR: *Medical Epidemiology*, 4th Edition: http://www.accessmedicine.com

Collectively work together to improve the health and wellness of people, places, and communities

References, Further Readings, and Additional Resources

- *Medical Epidemiology* (Greenberg, Daniels, Flanders, Eley, Boring III)
  - https://accessmedicine-mhmedical-com.proxy1.lib.tju.edu/book.aspx?bookid=337
- *Study Designs*
  - *https://www.cebm.net/wp-content/uploads/2014/06/CEBM-study-design-april-20131.pdf*
- *Incidence and Prevalence*
  - https://www.youtube.com/watch?v=1jzZe3ORdd8&feature=emb_title
- *Measures of Association and Significance*
  - http://critical-numbers.group.shef.ac.uk/lecturenotes/Manual.pdf
  - http://critical-numbers.group.shef.ac.uk/glossary/clinical_significance.html
  - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4640017/
  - http://www.pitt.edu/~bertsch/risk.pdf
  - https://www.cancerresearchuk.org/health-professional/cancer-statistics/cancer-stats-explained/our-calculations-explained#heading-Eleven
  - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938757/
  - https://www.slideshare.net/terryshaneyfelt7/what-does-an-odds-ratio-or-relative-risk-mean
  - https://sph.unc.edu/wp-content/uploads/sites/112/2015/07/nciph_ERIC11.pdf

References, Further Readings, and Additional Resources

- *Hazard Ratios*
  - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3932959/
  - https://www.statisticshowto.com/hazard-ratio/
- *P-Values*
  - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4111020/#:~:text=The%20width%20of%20the%20confidence,the%20reliability%20of%20the%20estimate
- *Evidence-Based Medicine*
  - https://libguides.ecu.edu/c.php?g=17486&p=97640
- *Evaluating and Assessing Literature*
  - https://lit.libguides.com/c.php?g=664499&p=4701470
- *Decision Analysis*
  - https://accessmedicine-mhmedical-com.proxy1.lib.tju.edu/content.aspx?bookid=337&sectionid=39810370#546354

Questions?

Thank you!

More questions, comments, or suggestions?
Email: Allison.Casola@jefferson.edu | Twitter: @arcasola

Jefferson

Thomas Jefferson University