Critical Inquiry

Adam Lutz, PT, DPT
PhD Student
University of South Carolina
Fall 2017 – Critical Inquiry

ATI PHYSICAL THERAPY

“Medicine is an evolving science. The accepted truths of a generation ago, or even last year, may have since been discredited.”

Consumer Reports on Health, 2000

ATI PHYSICAL THERAPY

What is Evidenced Based Medicine?

"The integration of individual clinical expertise with the best available clinical evidence from systematic research."

- David L Sackett, W Scott Richardson, William Rosenberg, R Brian Haynes *Evidence Based Medicine--How to Practice and Teach EBM*, 1996

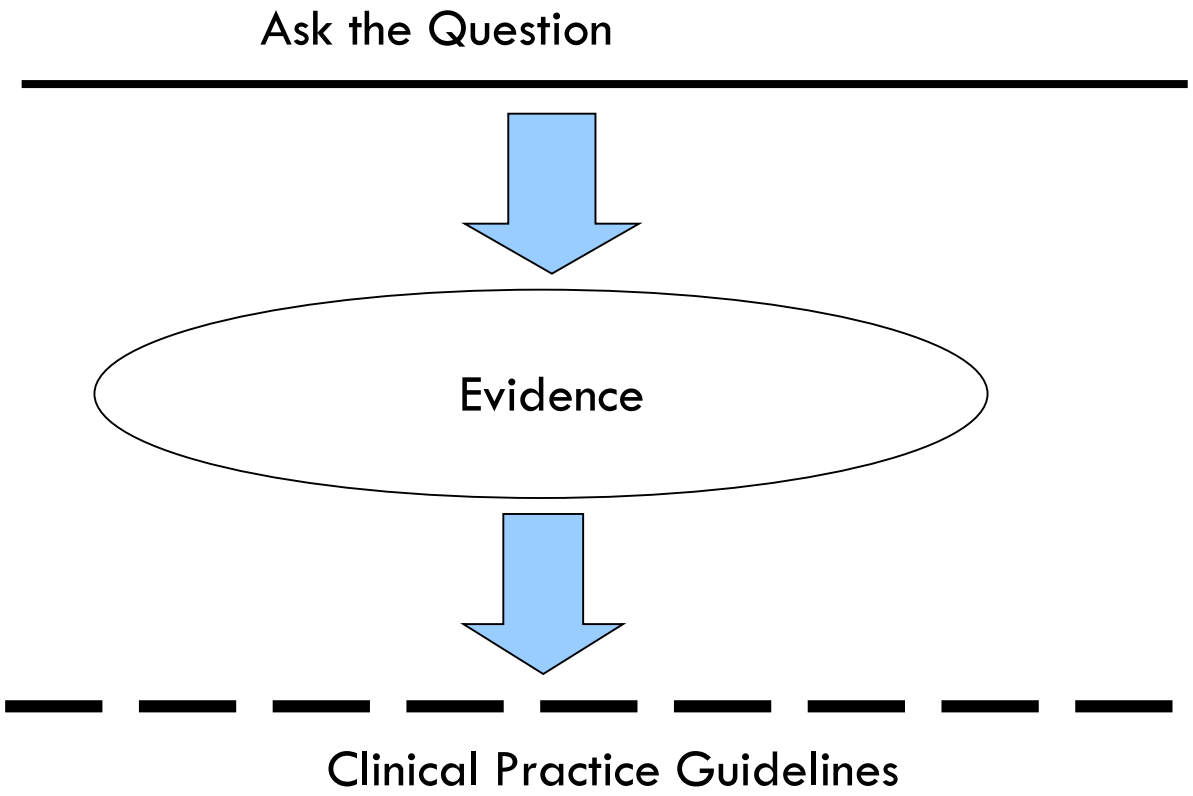
ATI PHYSICAL THERAPY

Dunning-Kruger Effect

ATI PHYSICAL THERAPY

Practicing Evidence-Based Medicine

- Ask the right question
- Assess the evidence
- Summarize the evidence
- Develop clinical guidelines
  - Which patient population
  - Which clinicians
- Recommendation
  - Useful
  - More evidence needed
  - Limitations



AHRQ Publication No. 02-E016 April 2002

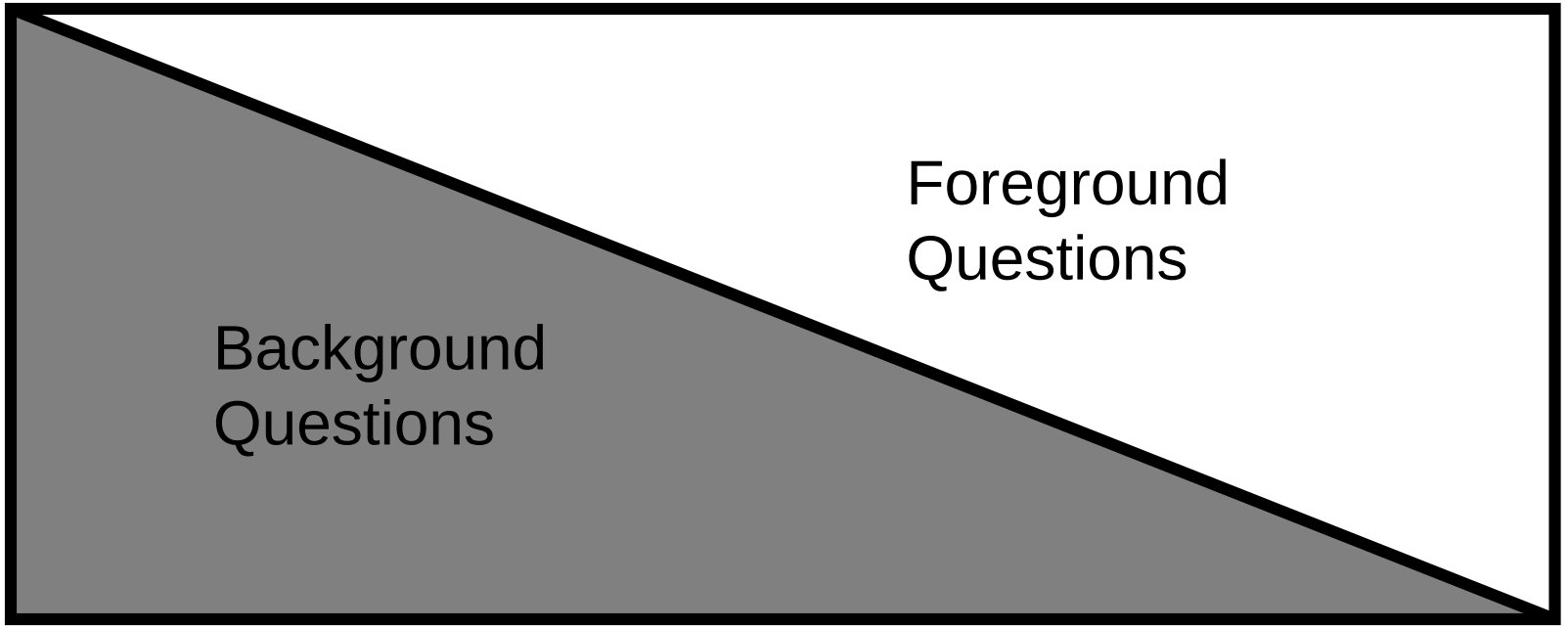
Ask Good Clinical Questions:

- **Background Questions**
  - **General** clinical questions for background knowledge.
  - Can be answered by using “background” resources, e.g., current textbooks, and narrative reviews

- **Foreground Questions**
  - **Specific** questions (Diagnosis, Etiology, Prognosis, Therapy) about your patient
  - Need latest resources
  - Answered by primary (original research) and secondary (systematic reviews and practice guidelines) sources

ATI PHYSICAL THERAPY

Effect of Clinical Experience



Quick Reference / Dictionary

- p – probability
- n – sample size
- s – standard deviation of a sample
- s^2 – variance of a sample
- α – pre-determined level of significance (typically .05)
  - Associated with Type I error (“false positive”)
- B – probability associated with type 2 error (typically .2)
  - Associated with Type II error (“false negative”)
  - 1- B = ‘power’ (typically .8)
- Power – probability that a test will lead to rejection of the null hypothesis
  - Function of α, s^2, n, and effect size (ES) of experimental variable

ATI PHYSICAL THERAPY

Quick Reference / Dictionary

- r – Pearson's coefficient of correlation
- R^2 – proportion of variance in the dependent variable that explained by the independent variables in the prediction equation
- $\bar{x}$ – sample mean
- t – test statistic for t-test
- F – test statistic for ANOVA
- (1-α) CI <or more commonly> 95% CI – confidence interval
  - Since α is typically .05, 1-α is typically .95 or 95%
- Σ – sigma; read 'the sum of'
  - May not see in literature as much but in equations to obtain most statistics
- χ^2 – chi square; test statistic of significance for nominal data
- df – degrees of freedom

ATI PHYSICAL THERAPY

Quick Reference / Dictionary

- Scales of Measurement
  - Categorical (qualitative)
    - Nominal
      - Descriptive in nature (yes, no; male, female; blood type)
      - Often coded 0/1
    - Ordinal
      - Rank-ordered
        - normal, good, fair

ATI PHYSICAL THERAPY

Quick Reference / Dictionary

- Scales of Measurement
  - Numeric (different structures for describing numeric data - red & dark grey)
    - Discrete
      - Numbers can only take on specific values
        - Number of ER visits at a hospital in an hour
        - Number of pregnancies for a patient or population
    - Continuous
      - Can take on any numeric value
    - Interval
      - Arbitrary zero (e.g. temperature in Fahrenheit or Celcius)
    - Ratio
      - Absolute zero (e.g. temperature in Kelvin…like we always use)

ATI PHYSICAL THERAPY

Quick Reference / Dictionary

Sources: Indiana.edu, science.psu.edu

- Confounding
  - When the relationship between exposure (IND var) and outcome (DEP var) is distorted by the presence of another variable (confounder). Confounding can be positive (observed association is biased away from the null) or negative (observed association is biased toward the null)

- Internal validity
  - How well an experiment is done, avoids confounding. The less chance for confounding in a study, the higher its internal validity is

ATI PHYSICAL THERAPY

Quick Reference / Dictionary

Sources: Indiana.edu, science.psu.edu

- External validity
  - How well data and theories from one setting apply to another
    - Bench science: will this finding be consistent in the ‘real world’?
    - Applied/ clinical science: can we assume that this finding would apply to MY PATIENTS?
- Internal-External Validity Trade-off?
  - If you have a ton of control over the population being studied, might be less generalizable

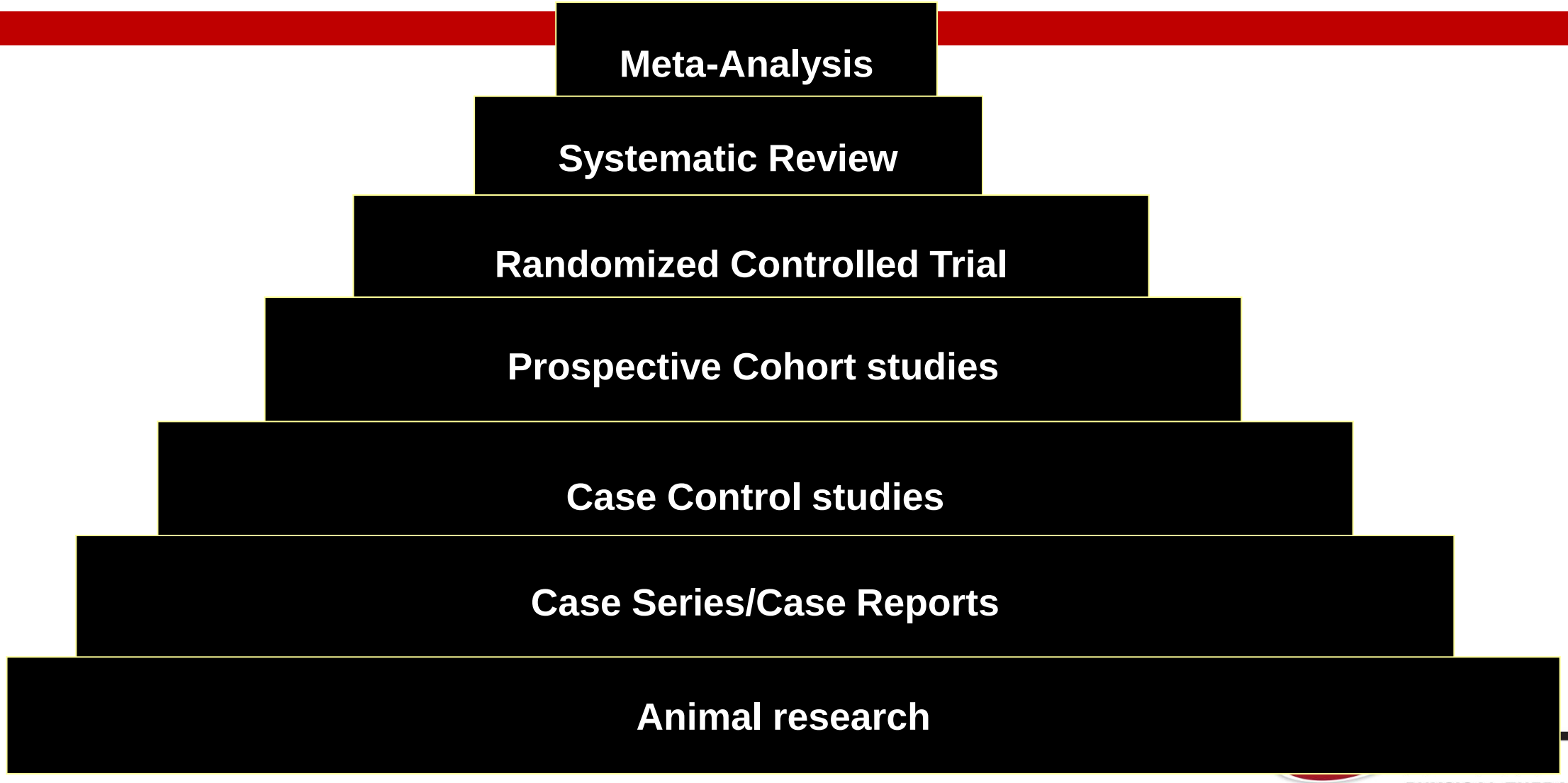
ATI PHYSICAL THERAPY

Variables

- Independent
  - ‘predictor’ variable
  - Condition, intervention, or characteristic that will predict or cause a given outcome
    - Examples: age, sex, race, income, zip code, …
- Dependent
  - Response or effect that is presumed to vary depending on the IND variable
    - Examples: pain, ROM, strength, change in outcome, …

ATI PHYSICAL THERAPY

Evidence Pyramid



PHYSICAL THERAPY

Animal Studies

Source: Foundations of Clinical Research; Portney, Watkins

- Basic/bench science; not clinical/application in nature
- Can be *Descriptive, *Exploratory, or *Experimental
  - Example in physical therapy might be looking at tissue response following a certain treatment application prior to use in humans
    - Post-operative management of stem-cell RTC repair in rats to assist in understanding of expected healing time in humans
  - Not common in our profession

ATI PHYSICAL THERAPY

Case Series/Reports

Source: Foundations of Clinical Research; Portney, Watkir

- Clinical/Applied Research
- *Descriptive Research
  - Description of interesting, new and unique case
  - In-depth description of an individual's condition or response to treatment
  - Can focus on group, institution, social unit (school, community, family)
  - When similar cases are reported, called 'case series'

ATI PHYSICAL THERAPY

Case-Control Studies

Source: Foundations of Clinical Research; Portney, Watkins

- Clinical/Applied Research
- Epidemiologic[x] / Observational Investigation
- *Descriptive or *Exploratory Research
  - Case-group identified to have condition to be studied
  - Control-group identified to NOT have condition to be studied
  - Retrospective in nature as we look back to determine differences in exposures or other factors that might contribute to condition
  - Typically cheaper than large prospective studies
  - Analysis: T-tests, chi-square (χ^2)

[x]epidemiology-a branch of medical science that deals with the incidence, distribution, and control of disease in a population
-Merriam-Webster

ATI PHYSICAL THERAPY

Cohort Studies

Source: Foundations of Clinical Research; Portney, Watkins

- Clinical/Applied Research
- *Descriptive or *Exploratory Research
  - A cohort is defined by *exposure
  - Study may include multiple groups (cohorts)
  - Cohorts (exposed vs unexposed) are followed (prospectively or retrospectively) to determine if they develop the condition/disorder/outcome
  - Subjects are interviewed or observed to determine the presence or absence of certain exposures, risks, or the natural history of condition
  - Analysis: T-tests, ANOVA's, ANCOVA's, chi-square (χ^2), multivariate linear/logistic regression

ATI PHYSICAL THERAPY

Randomized Control Trial (RCT)

Source: Foundations of Clinical Research; Portney, Watkins

- Clinical/Applied Research
- *Experimental Research
  - “Gold Standard”
  - Random allocation of ‘exposure’/intervention to experimental group; comparison against control group that is provided with placebo intervention
  - Generalizable—except, those that would allow for randomization into a surgical group or even a PLACEBO surgery are a special population themselves so not representative of everyone
  - Analysis: T-tests, ANOVA’s, ANCOVA’s, confidence intervals (typically 95%CI), multivariate linear/logistic regression

ATI PHYSICAL THERAPY

Quasi-Experimental Design

Source: Foundations of Clinical Research; Portney, Watkins

- Clinical/Applied Research
- *quasi-Experimental Research
  - In the ‘real world’ we often aren’t able to create an environment in which a study is entirely experimental in nature, so we control as much as we can and state the limitations of the study accordingly
    - Subject to concerns over internal validity due to inability to control for differences that exist between subjects that are grouped together as a cohort
    - May achieve greater generalizability (external validity)

ATI PHYSICAL THERAPY

Systematic Review

Source: Oxford School of Medicine

- Exploratory/Experimental Research
- High quality systematic reviews seek to:
  - Identify all relevant published and unpublished evidence
  - Select studies or reports for inclusion
  - Assess the quality of each study or report
  - Synthesize the findings from individual studies or reports in an unbiased way
  - Interpret the findings and present a balanced and impartial summary of the findings with due consideration of any flaws in the evidence
  - Can be qualitative or quantitative

ATI PHYSICAL THERAPY

Meta-Analysis

Source: Foundations of Clinical Research; Portney, Watkins

- *Experimental Research
- Secondary Analysis
  - Seeks previously published articles with similar designs/variables with the assumption that subjects are all part of a large, target population
  - Allows for single estimate of effect of an intervention on a large population vs. several effects, sometimes with opposing results

ATI PHYSICAL THERAPY

Statistics: T-test

Source: Foundations of Clinical Research; Portney, Watkins

- T-test
  - Paired-samples t-test
    - Pre-post testing – difference score
    - Twin studies
  - 2-sample/IND/unpaired t-test
    - 2 groups are compared
- Follows t distribution; uses t test statistic

ATI PHYSICAL THERAPY

Statistics: Chi-square χ^2

Source: Foundations of Clinical Research; Portney, Watkins

- Analysis of nominal data
- Compares observed frequencies and compares to expected frequencies

- Follows χ^2 distribution; uses χ^2 test statistic

ATI PHYSICAL THERAPY

Statistics: ANOVA

Source: Foundations of Clinical Research; Portney, Watkins

- Dependent variable is continuous; independent variable(s) ALL NOMINAL
- Used to describe the relationship between the continuous dependent variable and nominal independent variable

- Follows F Distribution; uses F test statistic

ATI PHYSICAL THERAPY

Statistics: ANCOVA (aka ANACOVA)

- Dependent variable is continuous; independent variables are a mix of continuous and nominal variables
- Describes the relationship between the continuous dependent variable and one or more nominal, controlling for the effects of one or more continuous variables

- Follows F Distribution; uses F test statistic

Source: Applied Regression Analysis and Other Multivariable Methods; Kleinbaum, Kupper, Nizam, Rosenburg

ATI PHYSICAL THERAPY

Statistics: Simple Linear/Logistic Regression

- Single independent ‘predictor’ variable for dependent variable, typically ‘Y’
  - Where
    - y-hat is predicted value
    - a is Y-intercept
    - b is slope
    - x is IND variable

$$\hat{y} = a + bx$$

- *Logistic is simply one of many ‘transformations’
  - Allows for manipulation of data
  - In biomedical/health research, we use the natural log although that is assumed and we just refer to it as ‘log transformation’

Source: Applied Regression Analysis and Other Multivariable Methods; Kleinbaum, Kupper, Nizam, Rosenburg

ATI PHYSICAL THERAPY

Statistics: Multiple Linear/Logistic Regression

- Multiple ‘factors’ involved
- Otherwise same basic principal as simple regression—just more x’s with more b’s
  - Example
    - Y-hat = $a + b_1x_1 + b_2x_2 + \ldots + b_kx_k$ where ‘k’ is the number of factors
- Can apply an interaction term
    - Y-hat = $a + b_1x_1 + b_2x_2 + b_3x_1x_2$; interaction is just the product of the 2 independent variables
- While a more complex model can provide a better prediction, sometimes researchers will use apply too many ‘factors’

Source: Applied Regression Analysis and Other Multivariable Methods; Kleinbaum, Kupper, Nizam, Rosenburg

ATI PHYSICAL THERAPY

What type of study??

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

http://www.hsl.unc.edu/lm/ebm/Supplements/QuestionSupplement.htm

ATI PHYSICAL THERAPY

An evaluation using experimental or quasi-experimental designs can be used by the clinician to determine the:

- **Efficacy** of an intervention – benefit of an intervention tested under controlled experiment conditions, usually with a control group
- **Effectiveness** of an intervention – benefit of an intervention as tested under “real world” conditions, often using quasi-experimental methods.

ATI PHYSICAL THERAPY

Participation - Who makes up my sample?

- Power analysis- provides basis to estimate *statistical significance*
  - Meaningful differences
    - *P values vs. MDC- MCID*

- Clinical significance is if you can actually say anything from your results….
  - How many patients/athletes present with given condition?
  - How many are likely to enroll?

ATI PHYSICAL THERAPY

“REAL WORLD”

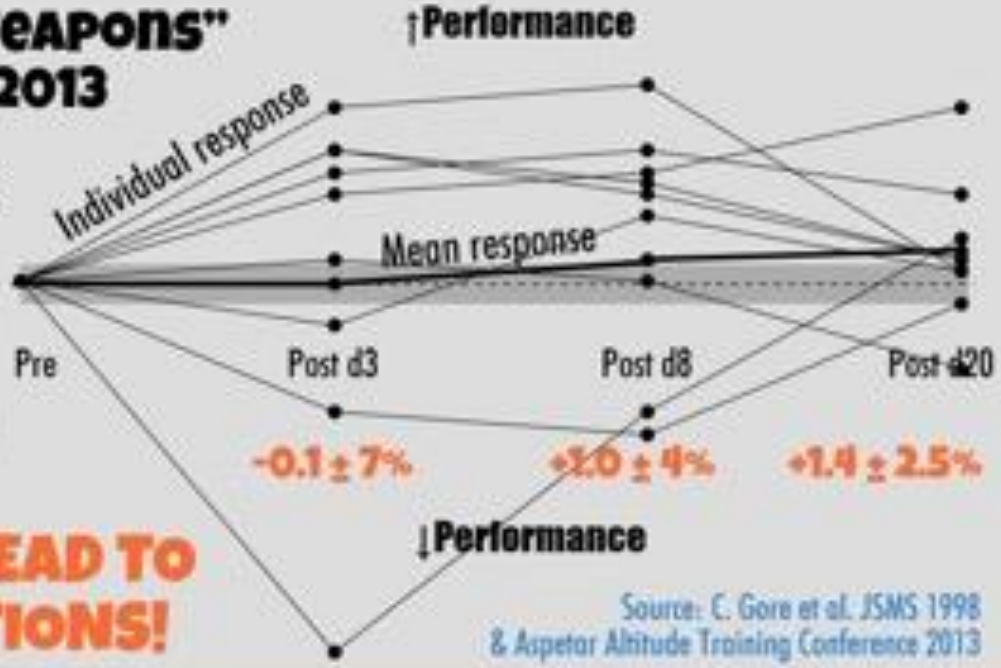
- All “real-world” treatment variation results from patients and providers making choices
- Large “real world” trials…who wants to be randomized?
- Who wants to adhere (or who DOES adhere)?

ATI PHYSICAL THERAPY

"STATISTICS ARE OUR WEAPONS" NICK BROAD,1974-2013

AN ILLUSTRATED EXAMPLE BY C. GORE

- 8 male cyclists
- All current or previous gold medallist and/or world record holder at senior or junior level
- Tested before and after a 31-days altitude training camp at 2690m (Pre, Post d3, Post d8, Post d20)



THE SAME DATASET CAN LEAD TO DIFFERENT INTERPRETATIONS!

Source: C. Gore et al. JSMS 1998
& Aspetar Altitude Training Conference 2013

P-VALUES

"COMPARED WITH THE BASELINE VALUE, THE GROUP MEAN PERFORMANCE WAS UNCHANGED ON ANY OF THE THREE MEASUREMENT DAYS AFTER THE CYCLISTS' RETURN FROM ALTITUDE"

Designed by @YLMSportScience

MAGNITUDE-BASED INFERENCES

Chances to observe & beneficial / trivial /harmful effect of altitude training

Post d3: 46/6/48%

Post d8: 61/9/30%

- Post d20: 79/10/11%

MANY COACHES WOULD BE HAPPY WITH A 79% CHANCE TO INDUCE A POSITIVE RESPONSE!

ATI PHYSICAL THERAPY

Critical Inquiry:
Clinical Application

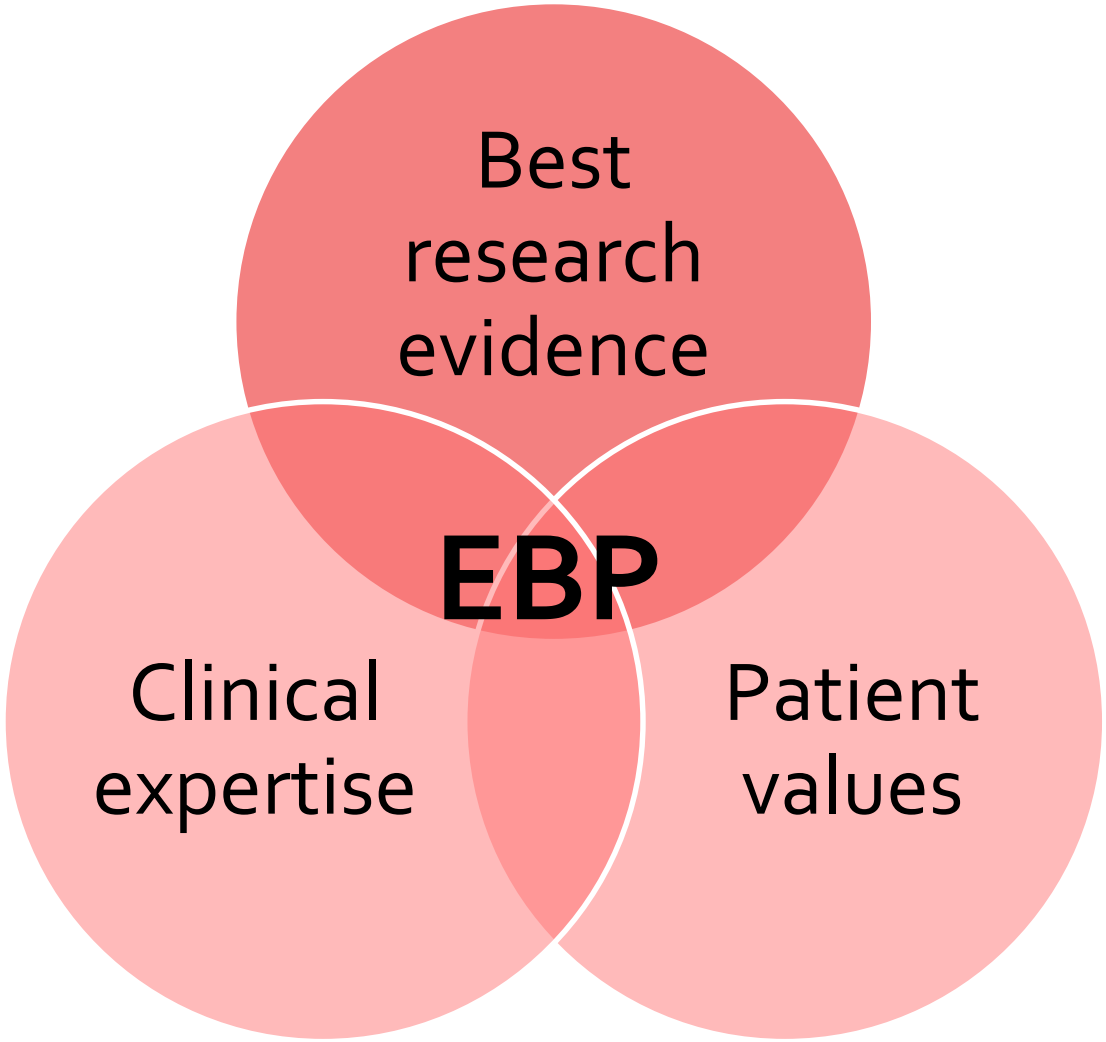
Courtney Chaaban, PT, DPT, SCS
ATI Physical Therapy
September 27, 2016

ATI PHYSICAL THERAPY

Objectives

- To improve confidence in:
  - Asking good questions
  - Reading and interpreting literature
  - Translating literature into clinical practice

Clinical Application for Critical Inquiry



Define your terms

Must ask a

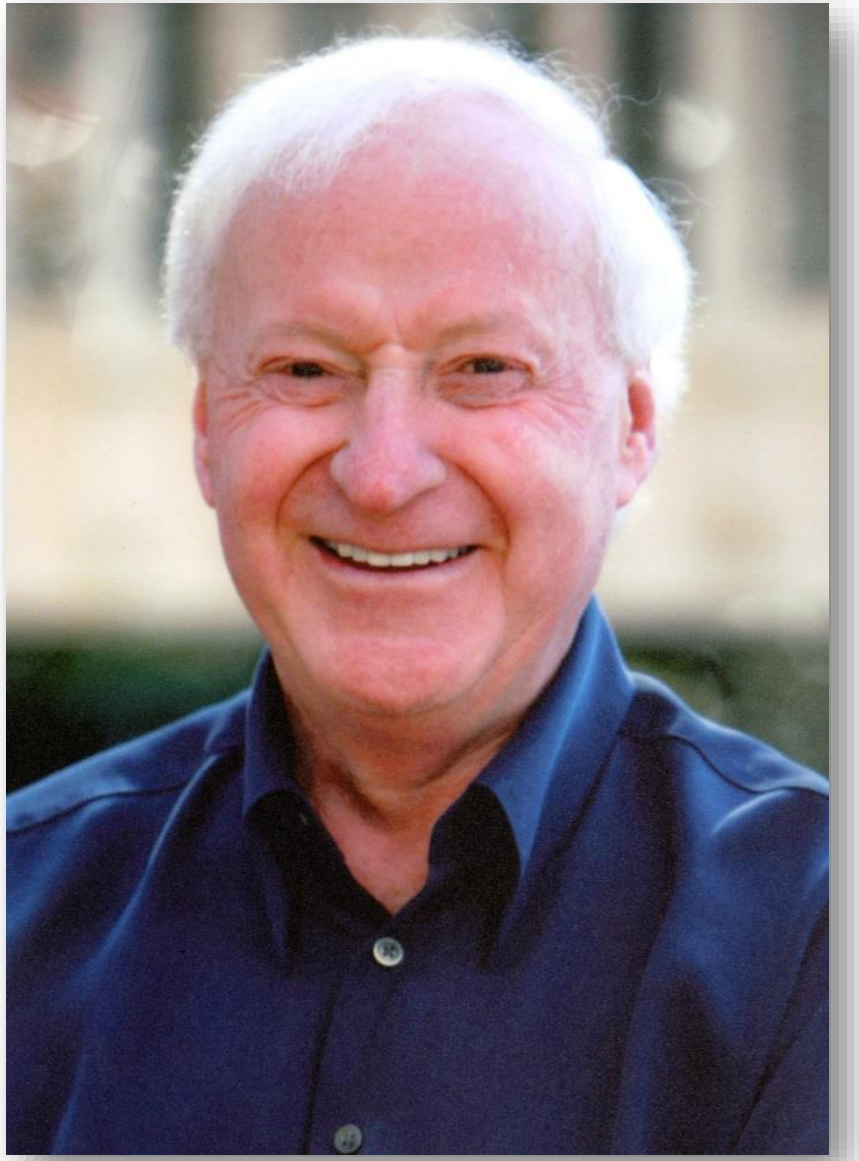
good question

to generate a

meaningful answer

Last week review:

- Comparison of means → t-test, ANOVA
- Association → correlation
- Prediction → regression



Outline

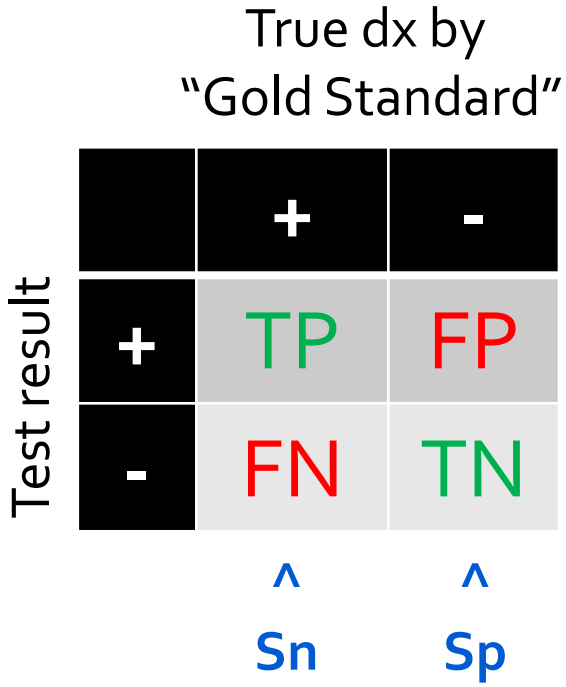
I. Validity
   I. Validity of discrete data
      - Sn, Sp, PPV, NPV, LR+, LR-
   II. Validity of continuous data
      - ROC curves
II. Clinical relevance
   I. Clinical relevance of discrete data
      - RRR, NNT
   II. Clinical relevance of continuous data
      - Effect size, MCID
III. Prognostic studies
   I. Survival analyses, RR, HR, OR

I. Validity of Discrete Data

True dx as determined by "Gold Standard"

Test result	Positive	Negative
Positive	True positive (TP)	False positive (FP) a.k.a. Type I Error
Negative	False negative (FN) a.k.a Type II Error	True negative (TN)

Validity of Discrete Data



Sensitivity = TP / (TP + FN)

- Proportion of people **with dx** who are correctly identified **positive** by test
- "True positives" – rarely negative when person has dx
- High sensitivity → helpful to **rule out** dx (low FN), good screen
  - Negative test: rule out diagnosis

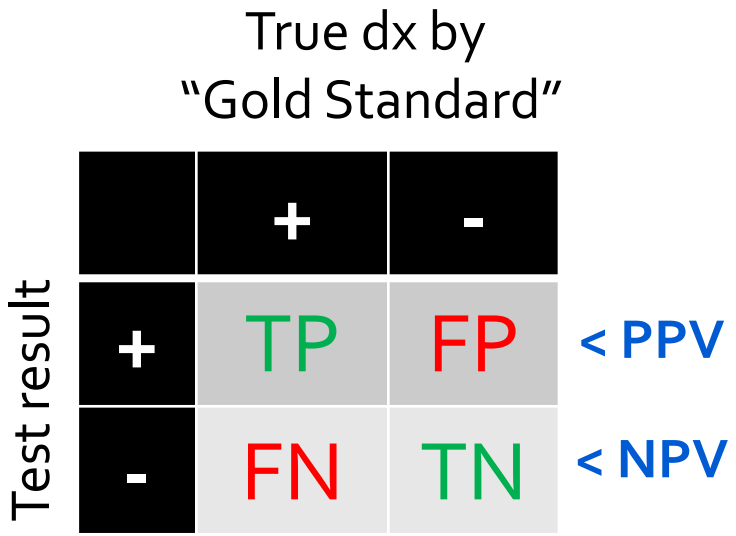
Specificity = TN / (TN + FP)

- Proportion of people **w/o dx** who are correctly identified **negative** by test
- "True negatives" – rarely positive when person doesn't have dx
- High specificity → helpful to **rule in** dx (low FP)
  - Positive test: rule in diagnosis

TEST RESULTS	WITH DISEASE	WITHOUT DISEASE
POSITIVE	TRUE POSITIVE (TP) = Have disease and have positive test	FALSE POSITIVE (FP) = No disease, but have positive test
NEGATIVE	FALSE NEGATIVE (FP) = Have disease, but have negative test	TRUE NEGATIVE (TN) = No disease and have negative test
	$\text{SENSITIVITY} = \frac{TP}{TP + FN}$	$\text{SPECIFICITY} = \frac{TN}{TN + FP}$

SpIN – higher specificity, fewer FP, more confidence 'ruling in'
SnOUT – higher sensitivity, fewer FN, more confidence 'ruling out'

Validity of Discrete Data



Positive predictive value (PPV) = TP / (TP + FP)

- Proportion of people identified **positive** by test **with dx**

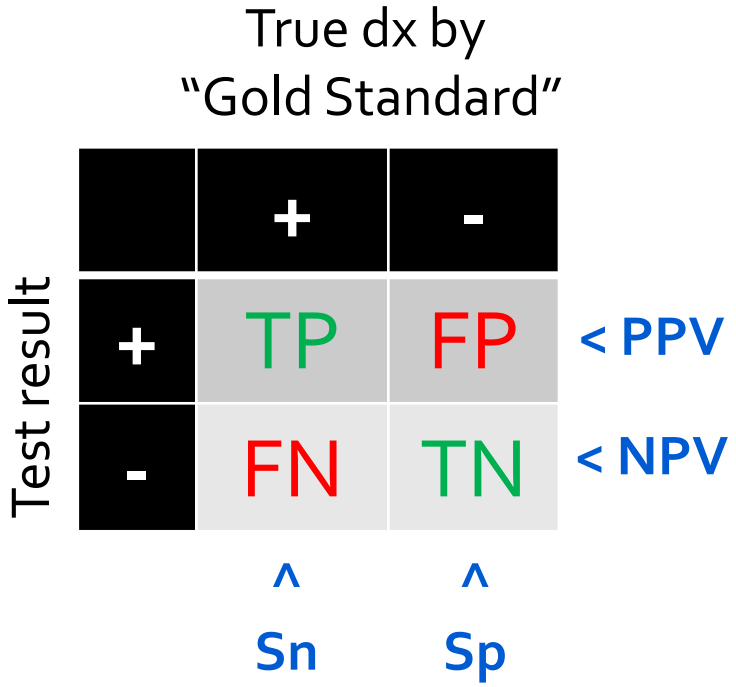
Negative predictive value (NPV) = TN / (TN + FN)

- Proportion of people identified **negative** by test **w/o dx**

PPV and NPV vary according to *prevalence*

- Affected by setting

Validity of Discrete Data



This is complicated. I want to know:

- The test is positive. *How much should this shift my suspicion the person has the condition?*
- The test is negative. *How much should this shift my suspicion the person has the condition?*

Likelihood ratios

- LR = $\frac{\text{probability of finding in pts w/ dx}}{\text{probability of same finding in pts w/o dx}}$

Validity of Discrete Data

True dx by "Gold Standard"

Test result	+	-	
+	TP	FP	< PPV
-	FN	TN	< NPV
	^ Sn	^ Sp	

Likelihood ratios

$$LR = \frac{TP\ rate}{TN\ rate}$$

Positive likelihood ratio (LR+) = Sn / (1 – Sp)

- Ratio of (+) test in people with & without dx
- LR+ > 1 → argues for dx

Negative likelihood ratio (LR-) = (1 – Sn) / Sp

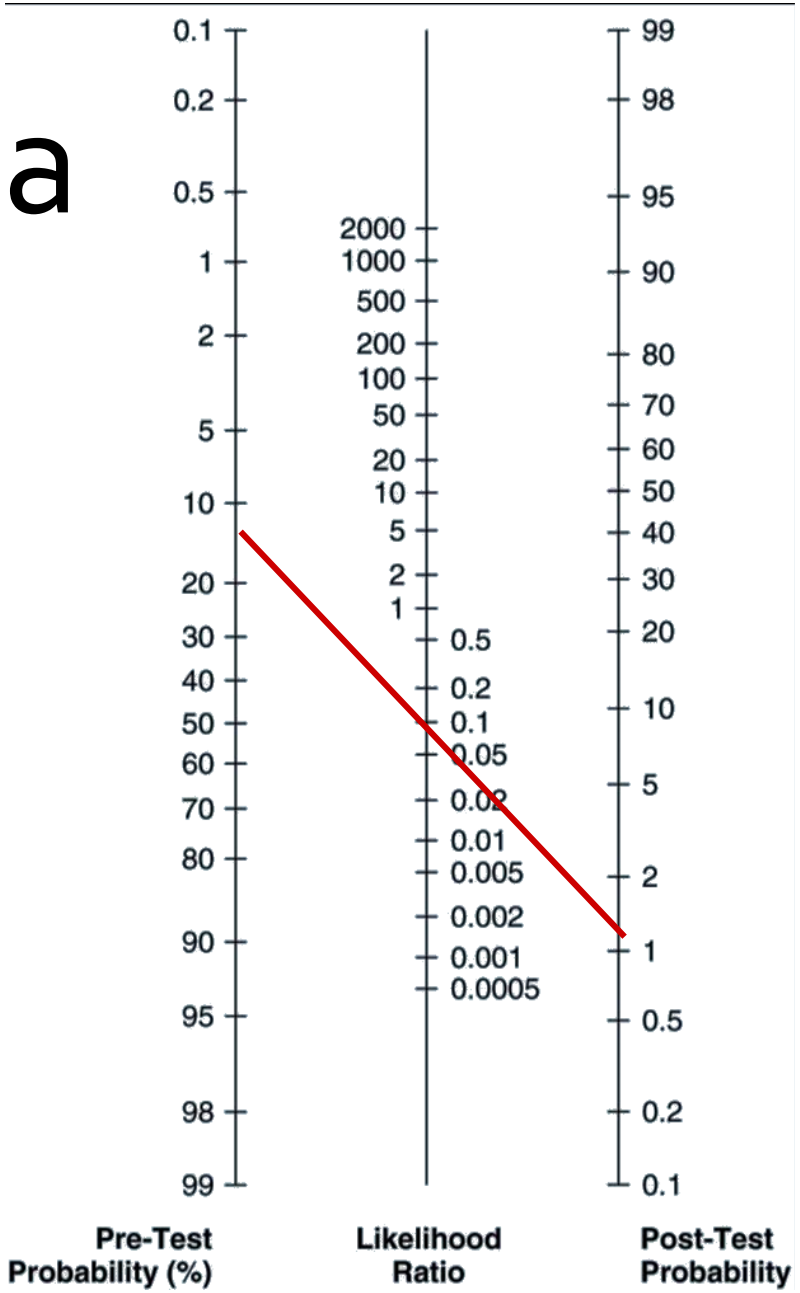
- Ratio of (-) test in people with & without dx
- LR- between 0 & 1 → argues against dx

Validity of Discrete Data

- Interpreting LRs: Using *nomograms* to estimate post-test probability
  - *Pre-test probability*: estimated by prevalence, what we know before testing
  - *LR*: what we know about the test
  - *Post-test probability*: what we think after interpreting test

Validity of Discrete Data

- Ex: Application of Ottawa ankle rules to acute ankle injuries
  - Pretest probability of ankle fracture = 15%
  - LR- = 0.08 (Bachman, BMJ, 2003)



What we think before the test

What we know about the test

What we think after the test

Validity of Discrete Data

Interpreting LRs:

Rough estimate:

2, 5, 10 →

15, 30, 45%

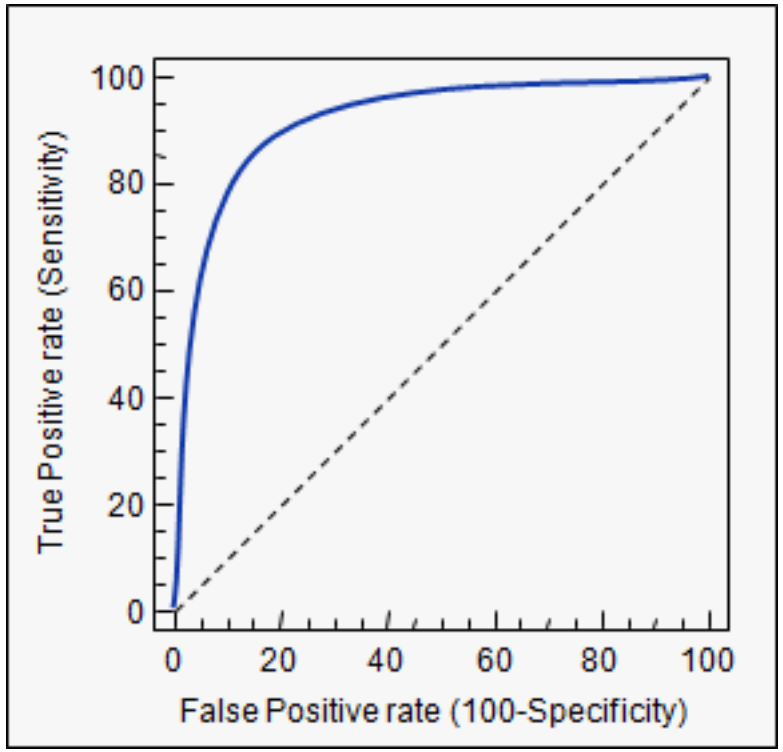
Likelihood Ratio	Approximate Change in Probability (%)*
Values between 0 and 1 decrease the probability of disease	
0.1	−45
0.2	−30
0.3	−25
0.4	−20
0.5	−15
1	0
Values greater than 1 increase the probability of disease	
2	+15
3	+20
4	+25
5	+30
6	+35
7	
8	+40
9	
10	+45

Validity of Discrete Data

- Interpreting LRs
  - *LR+ > 5* – increases your confidence to *rule in*
    - Likelihood increased 30% with positive finding
  - *LR- < 0.2* – increases your confidence to *rule out*
    - Likelihood decreases 30% with negative finding

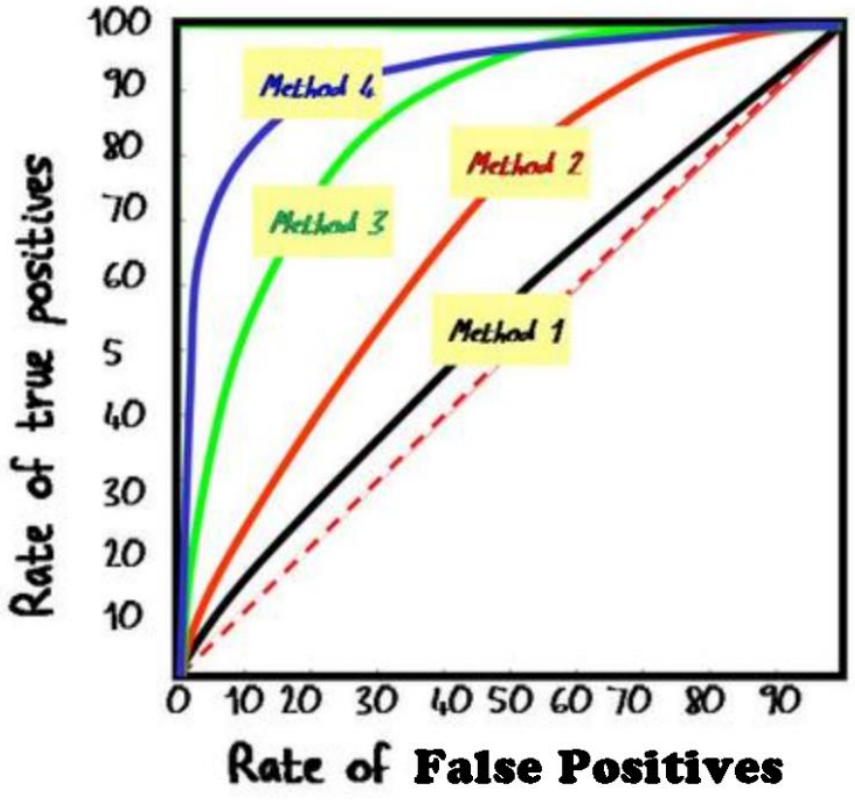
Validity of Continuous Data

- Receiver Operative Characteristic (ROC) curves
  - Attempts to answer: *What is maximum "correctness" of test? What is "cut score" for test?*
  - Meaning: area under curve (AUC) = probability of correct info from test



Validity of Continuous Data

- How to calculate:
  - "Cut score" is trade off between:
    - ↑ Sn (fewer missed ACLI risk)
    - ↑ Sp (fewer false ACLI risk)
  - Generally pick higher Sn (worse to miss ACLI risk)



II. Clinical Relevance of Discrete Data

- Want to know *magnitude* of effect (vs. count):
  - By *how much* does intervention ↓ risk of unwanted event, or ↑ risk of desirable event?
- Number needed to treat (NNT)
  - # of pts who must receive intervention to produce 1 positive outcome or avoid 1 adverse event (vs. control)

Clinical Relevance of Discrete Data

- Reducing risk (ex. ACL tear)
  - Absolute risk reduction (ARR)
    - ARR = % control with problem - % therapy group with problem
  - Relative risk reduction (RRR)
    - RRR = ARR / % control with problem
- NNT = 1/ARR

Clinical Relevance of Discrete Data

- Improving benefit
  - Absolute benefit increase (ABI)
    - ABI = % therapy group with outcome - % control group with outcome
  - Relative benefit increase (RBI)
    - RBI = ABI / % controls with outcome
- NNT = 1/ABI

Clinical Relevance of Continuous Data

Effect size –

- Attempts to answer:
  - *How effective is given treatment?*
  - Evaluates treatment effect independent of sample size
- How to calculate:
  - ES = change score / avg sd
- Meaning:
  - Expression of magnitude of difference between 2 sample means (see chart)

Interpretation

Value	Meaning
> 0.8	Big treatment effect
0.4 - 0.8	Confident test effect exists
0.2 – 0.4	Treatment effect small enough clinician may not perceive benefit
< 0.2	What treatment effect?

Clinical Relevance of Continuous Data

- Minimal Clinically Important Difference (MCID)
  - Attempts to answer:
    - *Is the effect meaningful to my pts?*
  - Meaning:
    - Value is smallest difference a pt would perceive as beneficial
  - How to calculate:
    - Outcome of interest vs. criterion outcome (GROC, etc)

III. Prognostic Studies

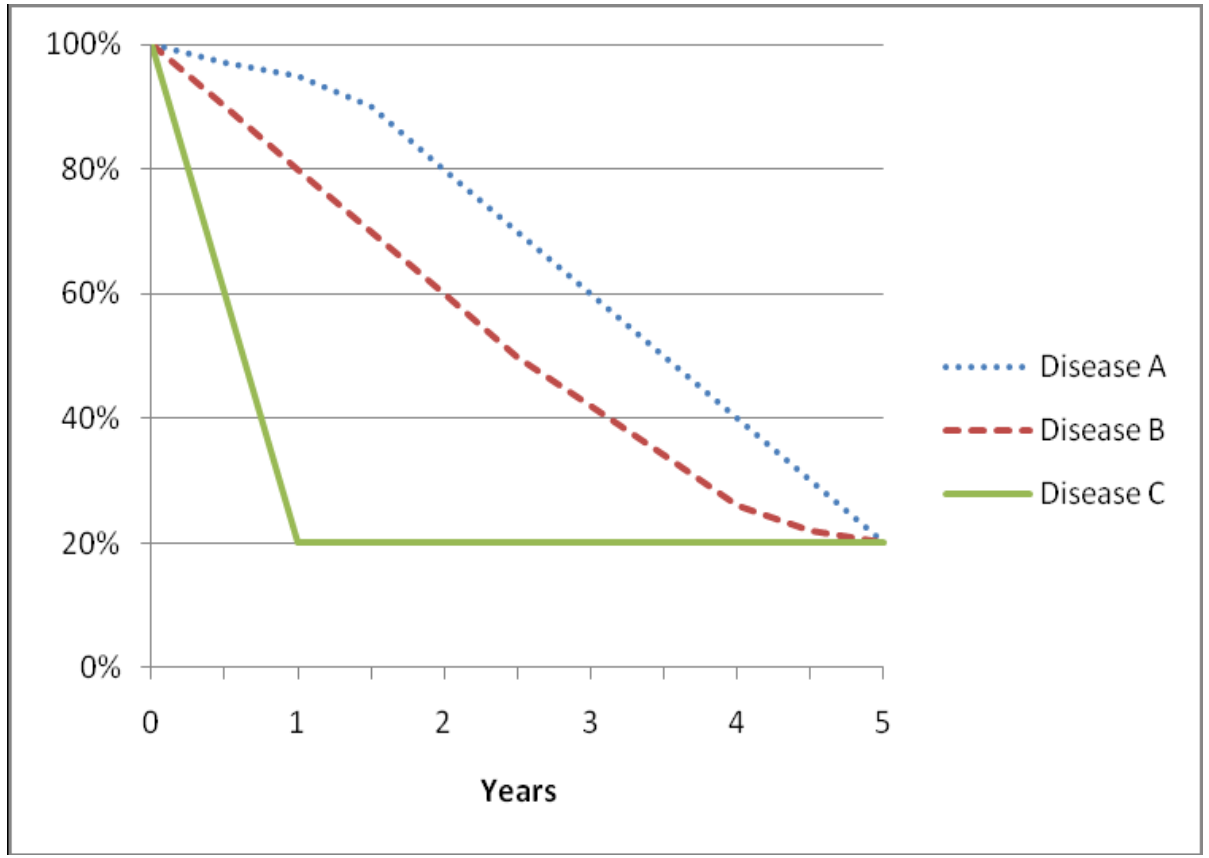
- Definition: *prognosis* is process of *predicting* the future about a pt's condition
- Elements of prognosis:
  - Outcomes possible
  - Likelihood outcomes will occur
  - Time frame required for achievement

Prognostic Studies

- Prognostic study results:
  - *Rates* (proportion experiencing event)
    - Relative risk: prospective, longitudinal estimate
    - Hazard ratio: prospective, instantaneous estimate
    - Odds ratio: retrospective, longitudinal estimate
  - *Survival curves* ("survival" of cohort over time)

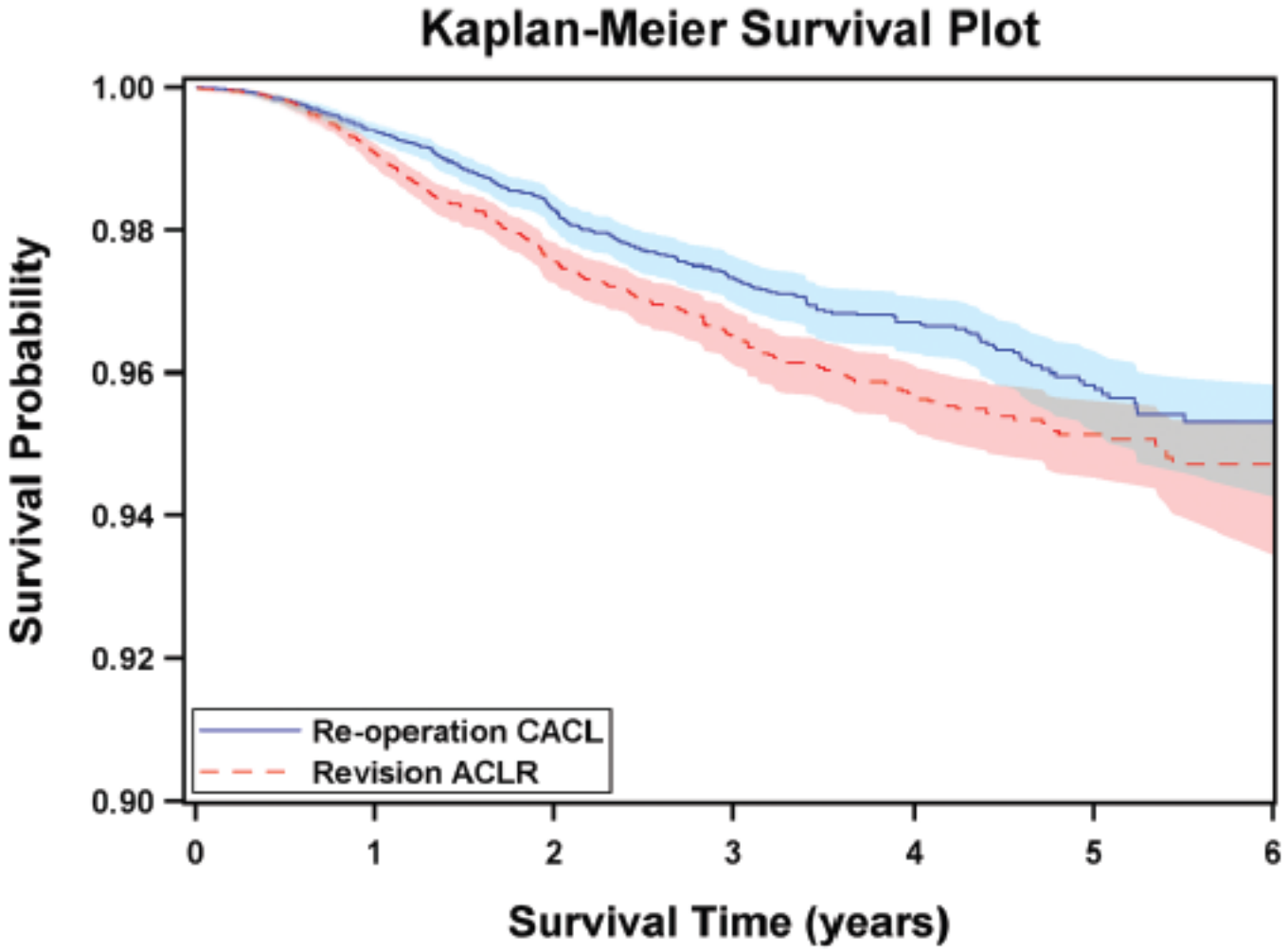
Prognostic Studies

- What survival curves add:



Prognostic studies

- Survival curves
  - Factor: intervention?
  - Outcome: rev. ACLR or CACLR



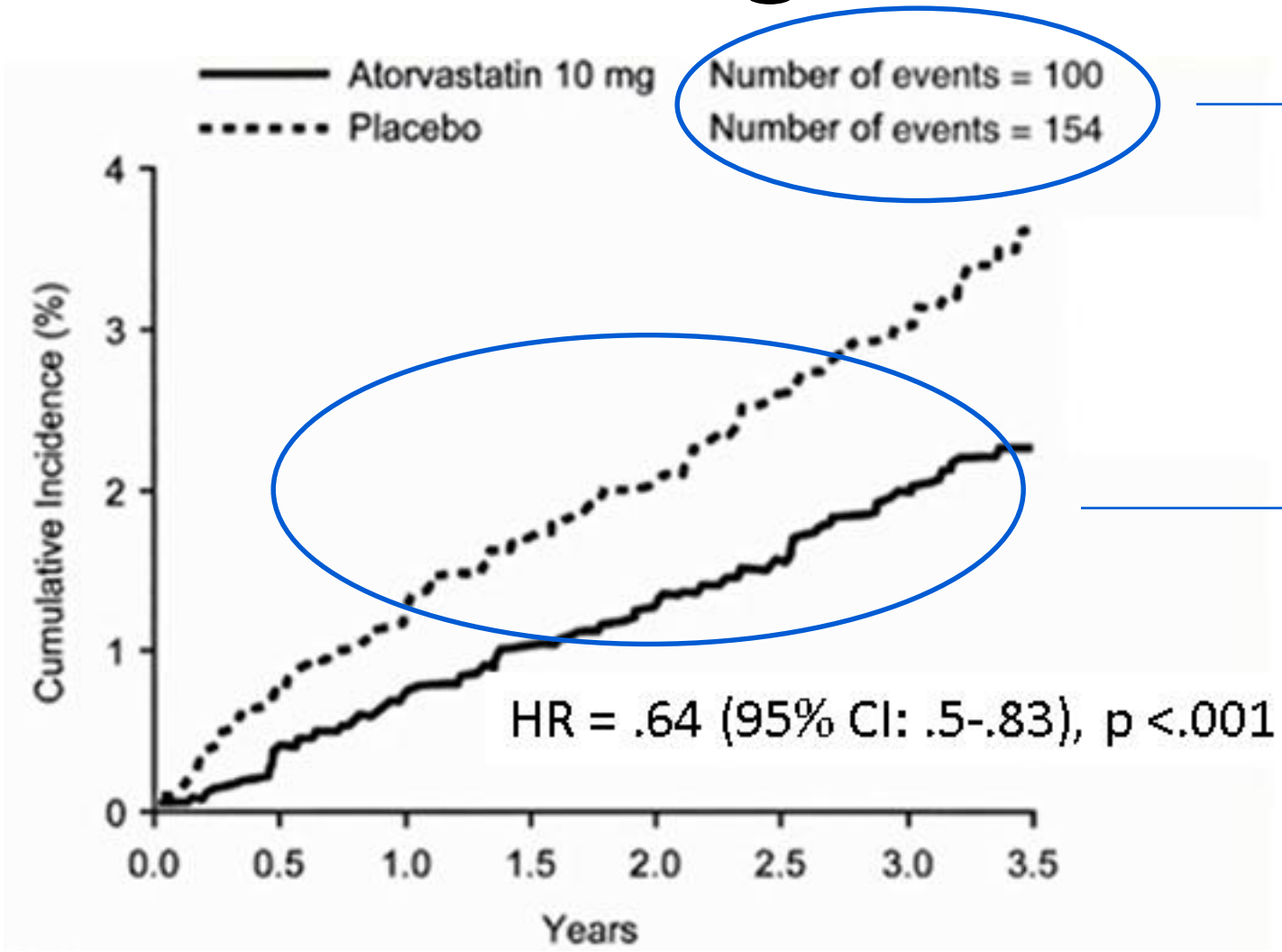
Prognostic Studies

- Relative risk
  - Likelihood pt "*exposed to a factor*" will develop outcome of interest
  - *Cumulative* over study with endpoint
  - Meaning: example value of 0.7 would mean intervention group is 0.7x as likely to develop outcome of interest

Prognostic Studies

- Hazard ratio
  - *Instantaneous* risk during study time period
  - Meaning: chance of event occurring at any time in intervention (vs control)

Prognostic Studies



Relative Risk is concerned with the total events over time

Hazard ratio is concerned with likelihood of events at any time

Prognostic Studies

- Odds ratio
  - Retrospective estimate of relative risk
    - Subjects already have outcome of interest (example – already have second ACLI)
  - Interpretation:
    - < 1 → decreased odds of outcome
    - > 1 → increased odds of outcome
  - Example: group is x times as likely to have outcome