Stats & Research Analysis
Critiquing the evidence

Tadd Hellwig, PharmD, BCPS, FASHP
Professor, South Dakota State University
College of Pharmacy and Allied Health Professions

Pharmacist Objectives

- Review common statistical analysis used when evaluating clinical trials
- State differences in statistical significance and clinical significance
- Define steps to complete when reviewing a clinical trial
- Calculate and define number needed to treat, relative risk, and relative risk reduction

Technician Objectives

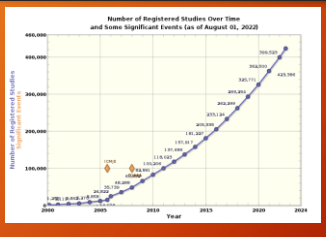
- Review importance of clinical trials
- Identify a significant p-value
- Recognize the difference in statistical significance and clinical significance
- Define the number needed to treat

Disclosure

- ***"I have had no financial relationship over the past 24 months with any commercial sponsor with a vested interest in this presentation"***

From all topic possibilities, why this topic???

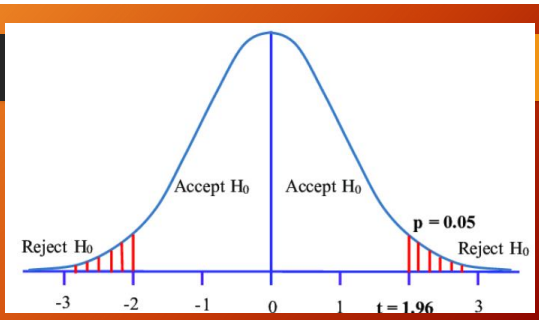
Where to start

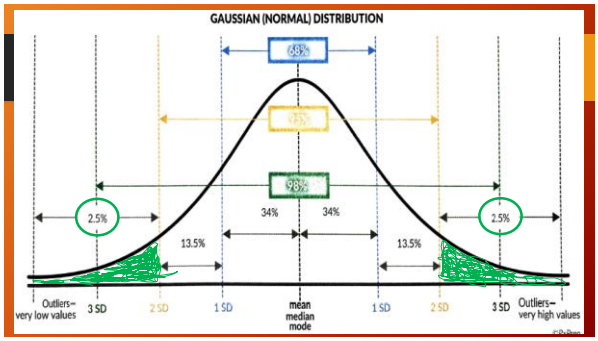


Quick Baseline Review- The P-value

- Null hypothesis (H_0)
  - Tmt 1 = Tmt 2
- Reject the H_0
  - Tmt 1 ≠ Tmt 2
  - (There IS a significant difference in groups)
- Fail to reject the H_0
  - Tmt 1 = Tmt 2
  - (No difference between groups)
- P-value
  - Probability of obtaining observed results under the assumption that the null hypothesis is true
  - Aim for a P-value < 0.05

p < 0.05 - statistically significant
p >/= 0.05 - not statistically significant

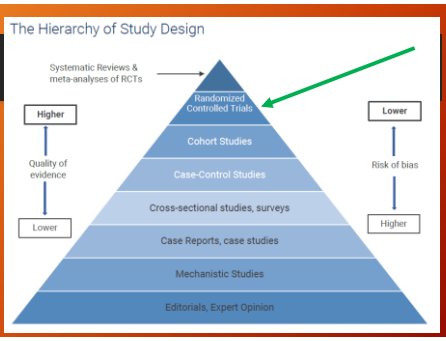




Types of Error

- **Type I error**
  - Null hypothesis was rejected in error
  - "False Positive"
  - Alpha or p-value helps prevent
- **Type II error**
  - Null hypothesis was accepted when it should habe been rejected
  - "False Negative"
  - Beta or power helps prevent

This is going nowhere fast!



Lets start with the abstract

- It is very easy to read the abstract!
- Separated out into the sections we are most curious about
  - Background
  - Methods
  - Results
  - Discussion??
  - Conclusions
- Never just do this and stop here!

ORIGINAL ARTICLE

Ticagrelor versus Aspirin in Acute Stroke or Transient Ischemic Attack

ABSTRACT

BACKGROUND

Ticagrelor may be a more effective antiplatelet therapy than aspirin for the prevention of recurrent stroke and cardiovascular events in patients with acute cerebral ischemia.

METHODS

We conducted an international double-blind, controlled trial in 674 centers in 33 countries, in which 13,199 patients with a nonminor ischemic stroke or high-risk transient ischemic attack who had not received intravenous or intracarterial thrombolysis and were not considered to have had a cardioembolic stroke were randomly assigned within 24 hours after symptom onset, in a 1:1 ratio, to receive either ticagrelor (180 mg loading dose on day 1 followed by 90 mg twice daily for days 2 through 90) or aspirin (300 mg on day 1 followed by 100 mg daily for days 2 through 90). The primary end point was the time to the occurrence of stroke, myocardial infarction, or death within 90 days.

RESULTS

During the 90 days of treatment, a primary end-point event occurred in 442 of the 6589 patients (6.7%) treated with ticagrelor, versus 497 of the 6610 patients (7.5%) treated with aspirin (hazard ratio, 0.89; 95% confidence interval [CI], 0.78 to 1.01; P=0.07). Ischemic stroke occurred in 383 patients (5.8%) treated with ticagrelor and in 441 patients (6.7%) treated with aspirin (hazard ratio, 0.87; 95% CI, 0.76 to 1.00). Major bleeding occurred in 0.5% of patients treated with ticagrelor and in 0.6% of patients treated with aspirin, intracranial hemorrhage in 0.2% and 0.3%, respectively, and fatal bleeding in 0.1% and 0.1%.

CONCLUSIONS

In our trial involving patients with acute ischemic stroke or transient ischemic attack, ticagrelor was not found to be superior to aspirin in reducing the rate of stroke, myocardial infarction, or death at 90 days. (Funded by AstraZeneca; ClinicalTrials.gov number, NCT01994720.)

Ticagrelor vs Aspirin in Stroke

- The primary endpoint occurred in:
  - Ticagrelor 442/6589 (6.7%)
  - Aspirin 497/6610 (7.5%)
  - HR 0.89 (95%CI 0.78 to 1.01; p= 0.07)

CONCLUSIONS

In our trial involving patients with acute ischemic stroke or transient ischemic attack, ticagrelor was not found to be superior to aspirin in reducing the rate of stroke, myocardial infarction, or death at 90 days. (Funded by AstraZeneca;

4. Ticagrelor is not recommended over aspirin for treatment of patients with minor acute stroke.	III: No Benefit	B-R	New recommendation.

What is the trial assessing

- What is the clinical trial design?
  - Hierarchy of clinical trials
- Is the patient population representative of all?
  - Inclusion/exclusion criteria
- What are the endpoints of the trial?
  - Many to choose from but will interpret them differently

Inclusion/Exclusion Criteria

Do you agree with these criteria for a clinical trial with a new drug for diabetes?

- **Inclusion Criteria**
  - T2DM
  - Age >18 years old
  - BMI <40
  - CVD diagnosis
  - A1c 7-9%
- **Exclusion Criteria**
  - Uncontrolled hyperglycemia
  - Liver disease
  - Cancer
  - Renal disease
  - Systemic steroids
  - Bariatric surgery

Baseline Characteristics

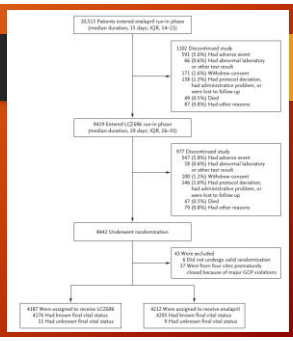
- Important to review
  - Do the patients represent the patients you see
  - How 'sick' were the patients
  - Were they receiving appropriate medications

TABLE 1. BASE-LINE CHARACTERISTICS OF THE PATIENTS.*

CHARACTERISTIC	PLACEBO GROUP (N=841)	SPIRONOLACTONE GROUP (N=822)
Age — yr	65±12	65±12
White race — %	86	87
Sex — no. (%)		
Male	614 (73)	603 (73)
Female	227 (27)	219 (27)
Blood pressure — mm Hg		
Systolic	122±20	123±21
Diastolic	75±11	75±12
Heart rate — beats/min	81±15	81±14
New York Heart Association class — no. (%)		
II	3 (0.4)	4 (0.5)
III	581 (69)	592 (72)
IV	257 (31)	226 (27)
Left ventricular ejection fraction — %†	25.2±6.8	25.6±6.7
Cause of heart failure — no. (%)‡		
Ischemic	453 (54)	454 (55)
Nonischemic	386 (46)	368 (45)
Medications — %		
Loop diuretics	100	100
ACE inhibitors	94	95
Digitalis	72	75
Aspirin	37	36
Potassium supplements	27	29
Beta-blockers	10	12

Run-In Periods

- Started with 10,513 patients with enalapril run-in phase
- Dropped to 9,419 patients during the Entresto run-in phase
- Total of 8,442 patients underwent randomization
  - Approx 80% of initial patients
- How does this affect your interpretation of the final results?



Why is clinical trial design important

Table 2. Components of the Primary and Secondary Efficacy Outcomes.*

Outcome	Cohort 1			
	Betrixaban (N=1914)	Enoxaparin (N=1956)	Relative Risk (95% CI)	P Value†
	no./total no. (%)			
Primary end point				
Primary efficacy outcome‡	132/1914 (6.9)	166/1956 (8.5)	0.81 (0.65–1.00)	0.054

CONCLUSIONS

Among acutely ill medical patients with an elevated D-dimer level, there was no significant difference between extended-duration betrixaban and a standard regimen of enoxaparin in the prespecified primary efficacy outcome. However,

Cohort 2				Overall Population			
Betrixaban (N=2842)	Enoxaparin (N=2893)	Relative Risk (95% CI)	P Value†	Betrixaban (N=3112)	Enoxaparin (N=3174)	Relative Risk (95% CI)	P Value†
no./total no. (%)				no./total no. (%)			
160/2842 (5.6)	204/2893 (7.1)	0.80 (0.66–0.98)	0.03	165/3112 (5.3)	223/3174 (7.0)	0.76 (0.63–0.92)	0.006

Endpoint here, Endpoint there, Endpoints everywhere

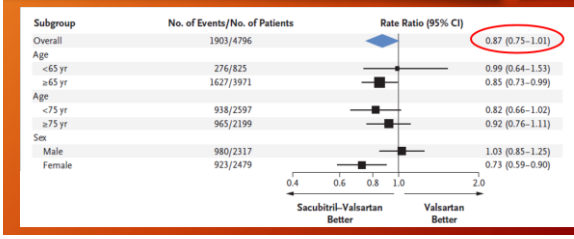
- **Primary**
  - Focus of the study
  - Usually what is powered for significance
- **Secondary**
  - Usually related to primary could be a surrogate of the primary
- **Clinical**
  - Occurrence of disease or event
- **Biomarker**
  - Nonclinical objective lab values
- **Surrogate**
  - Indicator or sign pointed towards a clinical endpoint
  - Benefits: attain these sooner!
- **Composite**
  - Combines multiple endpoints into a single endpoint
- **Exploratory**
  - Hypothesis generating, unknown true utility
- **Cancer oh my…**
  - Disease-free progression, Progression-free survival, Complete response, Overall response

How are they assessing the endpoint

- Superiority
  - Designed to detect significant differences between treatments
- Non-inferiority margin
  - Designed to detect if a new therapy is not worse than a previous therapy
  - Hope it is not unacceptably less efficacious than previous therapy
- Confidence intervals
  - Range of likely values of the endpoint in which you are 95% confident the true value resides

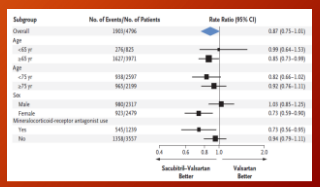
OR or HR < 1: event rate is lower in the treatment group than in the control group
OR or HR = 1: event rate is the same in the treatment and control arms. No advantage to the treatment
OR or HR > 1: event rate is higher in the treatment group than in the control group

Forest Plots

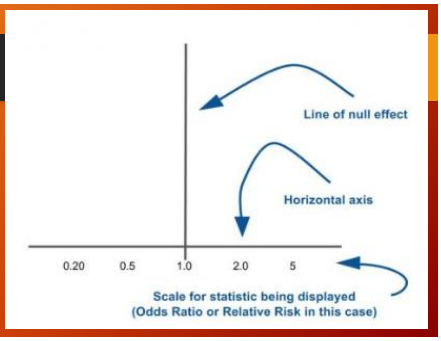


How to Interpret Forest Plots

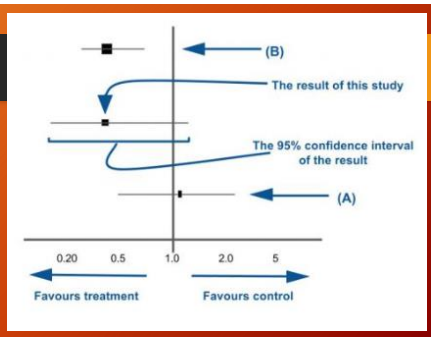
1. What are the 'axis' telling us
2. What are the 'lines' telling us



- Line of null effect will help determine statistical significance
  - (P-value)

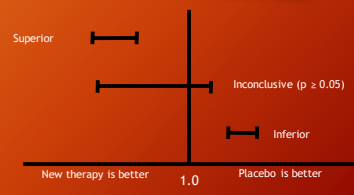


- Black boxes are the point estimates
- Lines represent the 95% CI
- Width of lines typically indicate sample size
  - Narrow = big sample size, more reliable
  - Wide = small sample size, less reliable



Forest Plot Interpretation

- Does the CI cross the null effect line?
  - 1.0 or 0
- YES
  - Then there is NO significant differences between groups
  - P ≥ 0.05
- NO
  - Then there IS a significant difference between groups
  - P < 0.05



Non-Inferiority trials

- Superiority trials
  - Goal: determine if one therapy is significantly better than another
- Non-inferiority trials
  - Goal: determine if one therapy is no worse than another by a pre-determined margin
- Why do a non-inferiority trial
  - Unethical to use placebo
  - Comparing to standard of care
- Non-inferiority margin
  - Predetermined margin of difference between 2 groups that is considered acceptable/tolerable for the new treatment to be considered 'similar' or 'not worse'

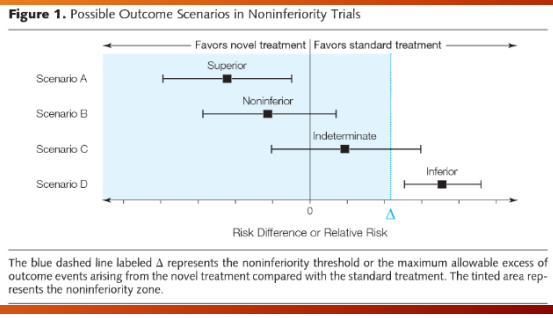
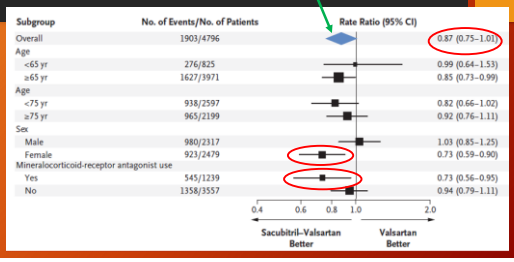


Figure 1. Possible Outcome Scenarios in Noninferiority Trials

The blue dashed line labeled Δ represents the noninferiority threshold or the maximum allowable excess of outcome events arising from the novel treatment compared with the standard treatment. The tinted area represents the noninferiority zone.

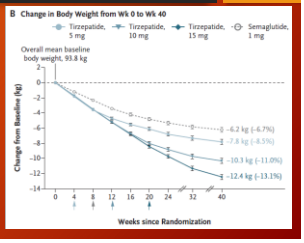
Forest Plots- Subsets



Endpoint

What type of endpoint is displayed in this trial?

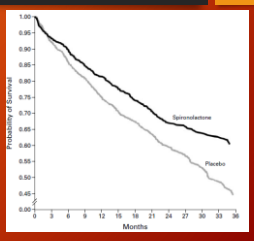
A. Clinical
B. Surrogate
C. Composite
D. Exploratory



Endpoint

What type of endpoint is displayed in this trial?

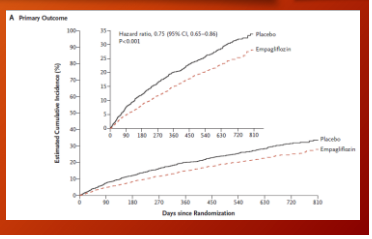
A. Clinical
B. Surrogate
C. Composite
D. Exploratory



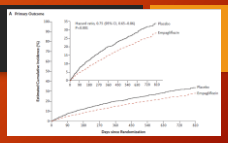
Endpoint

What type of endpoint is displayed in this trial?

A. Clinical
B. Surrogate
C. Composite
D. Exploratory



- The primary outcome was a COMPOSITE of cardiovascular death or hospitalization for heart failure

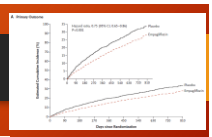


CONCLUSIONS

Among patients receiving recommended therapy for heart failure, those in the empagliflozin group had a lower risk of cardiovascular death or hospitalization for heart failure than those in the placebo group, regardless of the presence or absence of diabetes. (Funded by Boehringer Ingelheim and Eli Lilly; EMPEROR-Reduced

Where are the P-values???

Composite endpoints



Primary composite outcome — no. (%)	0.75 (0.65 to 0.86)
Hospitalization for heart failure	0.69 (0.59 to 0.81)
Cardiovascular death	0.92 (0.75 to 1.12)

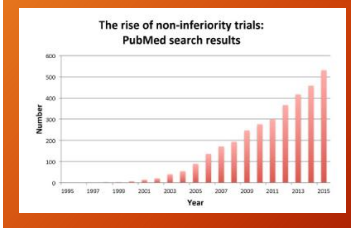
Variable	Empagliflozin (N = 1863)		Placebo (N = 1867)		Hazard Ratio or Absolute Difference (95% CI)†	P Value
		events/100 patient-yr		events/100 patient-yr		
Primary composite outcome — no. (%)	361 (19.4)	15.8	462 (24.7)	21.0	0.75 (0.65 to 0.86)	<0.001
Hospitalization for heart failure	246 (13.2)	10.7	342 (18.3)	15.5	0.69 (0.59 to 0.81)	
Cardiovascular death	187 (10.0)	7.6	202 (10.8)	8.1	0.92 (0.75 to 1.12)	

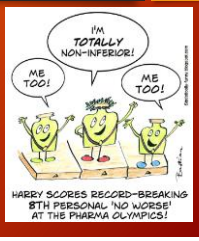
Graphical Representation

Primary composite outcome — no. (%)	0.75 (0.65 to 0.86)
Hospitalization for heart failure	0.69 (0.59 to 0.81)
Cardiovascular death	0.92 (0.75 to 1.12)

- What is the P-value for the Composite outcome?
- What is the P-value for Hospitalizations?
- What is the P-value for CV Death?

Non-inferiority Trials





Non-Inferiority Trials- Rivaroxaban

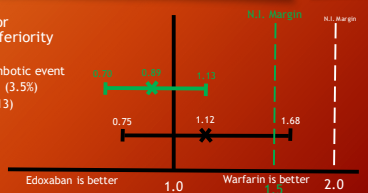
- Rivaroxaban vs. warfarin for treatment of pulmonary embolism- non-inferiority trial
  - Results for recurrent thrombotic event
  - Rivaroxaban (2.1%) vs warfarin (1.8%)
  - HR 1.12 (95% CI 0.75 to 1.68)
- What additional piece of information is needed?

Non-Inferiority Trials- Rivaroxaban

- Rivaroxaban vs. warfarin for treatment of pulmonary embolism- non-inferiority trial
  - Results for recurrent thrombotic event
  - Rivaroxaban (2.1%) vs warfarin (1.8%)
  - HR 1.12 (95% CI 0.75 to 1.68)
- Is this trial significant for non-inferiority?
- Do you agree with the N.I. margin?

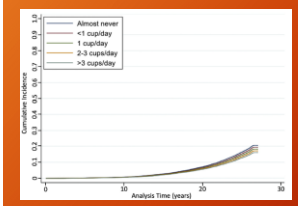
Non-Inferiority Trials- Edoxaban

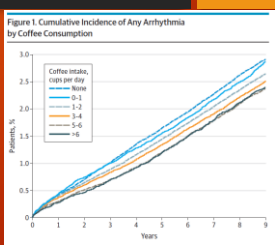
- EDOXABAN vs. warfarin for treatment of VTE: non-inferiority trial
  - Results for recurrent thrombotic event
  - Edoxaban (3.2) vs warfarin (3.5%)
  - HR 0.89 (95% CI 0.70 to 1.13)
  - N.I. Margin is 1.5
- Is this trial significant for non-inferiority?



Evaluate the graphs & tables

- Zoomed in graphs vs zoomed out





Don't just read the headlines

Which medication would you rather have?

A. Survival of 92.5% in treatment group vs. 90% in control group
B. Treatment with a new drug led to a 25% RRR in mortality
C. Treatment with a new drug led to a 2.5% reduction in mortality
D. A new drug will avoid 1 death in every 40 patients treated

Quick baseline review

- Absolute risk reduction (ARR)
  - Arithmetic difference between the event rates between groups
- Relative risk (RR)
  - Probability of an outcome in the exposed group to the probability of an outcome in an unexposed group
- Relative risk reduction (RRR)
  - Relative decrease in the risk of an event in the exposed group compared to the unexposed group
- Number needed to treat (NNT)
  - Number of people you need to treat to prevent one event

$$\text{Relative risk} = \frac{\text{Event rate in intervention group}}{\text{Event rate in control group}}$$

$$\text{Relative risk reduction} = 1 - \text{relative risk}$$

or

$$\text{Relative risk reduction} = \frac{\text{Absolute risk reduction}}{\text{Event rate in control group}}$$

$$\text{Absolute risk reduction (also know as "risk difference")} = \text{Event rate in intervention group} - \text{Event rate in control group}$$

$$\text{Number needed to treat} = \frac{1}{\text{Absolute risk reduction}}$$

Don't just read the headlines

	Trial #1 Treatment: 30% mortality Control: 40% mortality			Trial #2 Treatment: 7.5% mortality Control: 10% mortality	
RR	30%/40%	0.75			
RRR	1 - 0.75	25%			
ARR	40% - 30%	10%			
NNT	1 / 10%	10			

Don't just read the headlines

	Trial #1 Treatment: 30% mortality Control: 40% mortality			Trial #2 Treatment: 7.5% mortality Control: 10% mortality	
RR	30%/40%	0.75		7.5%/10%	0.75
RRR	1 - 0.75	25%		1 - 0.75	25%
ARR	40% - 30%	10%		10% - 7.5%	2.5%
NNT	1 / 10%	10		1 / 2.5%	40

Difference in RRR vs ARR

- The RRR measures our risk of something compared to something else
- ARR measures our risk of something compared to something else ……
  while accounting for the actual likelihood the event will happen in the first place

A RRR of 75% is the same thing of going from 40% to 10% in one trial compared to going from 1% to 0.25% in another trial but the ARR is drastically different

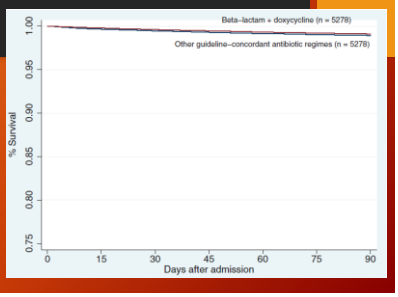
Clinical vs. statistical significance

- Objective: clopidogrel vs. aspirin to reduce recurrent stroke, MI, and death
- Primary outcome: recurrent ischemic event rates
  - Clopidogrel 5.33%
  - Aspirin 5.83%
  - (p = 0.043)

- Is this result STATISTICALLY significant?
- Is this result CLINICALLY significant?

ARR= 5.83% -5.33% (ARR = 0.5%)
NNT= 1/0.5% (NNT = 200)

- Doxycycline vs. azithromycin in community acquired pneumonia
  - 10,000 patients
  - P< 0.001
- Is this result clinically significant?
  - Doxy: 98%
  - Azith: 97.3%



How best to stay Up-to-Date with trials

- Cheat! Have others do the work for you
  - Professional organizations
    - APhA, ASHP, ACCP, NCPA, others
      - Listserves
    - Sign up for Electronic table of contents
    - Journal Watch
    - Twitter
- Look for editorials and/or commentaries
- Always someone smarter than you, learn from them

Take Home Points

- Skim the abstract
- What is the trial design
- Evaluate the baseline characteristics
- What are the endpoints
- Are the results significant
- Interpret the graphs
- Read the discussion
- How to incorporate into practice

It's Finally Over

Pharmacist post-test question 1

Which of the following is the most accurate statement when evaluating clinical trials?

A. The abstract always provides an accurate assessment of the entire clinical trial
B. The trial's inclusion/exclusion criteria aren't as important when determining therapy for actual patients you treat
C. All endpoints should be assessed as having equal weight
D. A surrogate endpoint is not as strong as a clinical endpoint

Pharmacist post-test question 2

Which of the following most likely signifies BOTH a statistically significant result and a clinically significant result?

A. Reduction in LDL with a new treatment of 54 mg/dL vs placebo reduction of 25 mg/dL (p-value of 0.062)
B. Reduction of systolic BP with a new treatment of 7 mmHg vs HCTZ reduction of 4 mmHg (p-value of 0.003)
C. Reduction in mortality with a new CFH medication of 3.4% vs placebo reduction of 6.4% (p-value of 0.085)
D. Reduction in readmissions for CHF exacerbation with a new medication of 14.5% vs placebo of 24.5% (p-value of 0.015)

Pharmacist post-test question 3

Which of the following offers the more significant reduction in mortality of a new therapy?

A. Absolute risk reduction of 3%
B. Relative risk reduction of 33%
C. I need more information to decide

Technician post-test question 1

Which of the following demonstrates a significant p-value?

A. Less than 0.05
B. Less than 0.5
C. Less than 1.0

Technician post-test question 2

Which of the following would represent a more statistically and clinically significant reduction?

A. A reduction in pain symptoms
B. A reduction in mortality
C. A reduction in potassium levels
D. A reduction in cholesterol

Technician post-test question 3

The number needed to treat is the number of patients that you need to treat to prevent on additional bad outcome?

A. True
B. False