High-Powered Medicine

Landmark Clinical Trial Reviews

"Good doctors use both individual clinical expertise and the best available external evidence, and neither alone is enough." -*David Sackett, MD*

www.highpoweredmedicine.com

High-Powered Medicine

Landmark Clinical Trial Reviews

Quick Review of Biostatistics

Alex Poppen
Doctor of Pharmacy

High-Powered Medicine

Landmark Clinical Trial Reviews

Please Note

This is **not** meant to be a comprehensive review of biostatistics.

It is only meant to serve as a quick reference for analyzing and interpreting trial data.

Power

The ability of a trial to detect a statistically significant difference between treatment groups *when a difference truly exists*

- Higher power = higher quality

Power

Example

90% power means that there is a 90% probability that the difference between two groups is **NOT** due to chance

- It also means that there is a 10% chance for showing no difference between groups *when a difference truly exists*
    - AKA - false negative

Power

If power is set but **NOT** met and there is **NO** significant difference between treatment groups the results should be considered *inconclusive*

- Trial lacked power to confirm that the results were **not** due to chance alone

If power is not met but a *statistically significant difference is observed*, this is less of a concern

- Significant difference detected despite low power

P-value

The probability that the observed difference between two groups is due to chance alone

P-value

Power:
The ability of a trial to detect a statistically significant difference between treatment groups *when a difference truly exists*

P-value:
The probability that the observed difference between two groups for a *specific outcome* is due to chance alone

P-value

The **level of significance** (aka alpha) is the probability the trial investigators are willing to take that the results were due to *chance alone* (typically set at 0.05)

- Alpha = 0.05 = 5% odds for false positive

If the p-value is less than alpha then the difference between the two groups is considered *statistically significant*

Hazard Ratio

The probability of an event occurring in the active group compared to the control group

- HR < 1.00 = lower probability
- HR > 1.00 = higher probability
- HR 1.00 = no treatment difference

Confidence Interval

The range of values in which the *true value* for an outcome resides

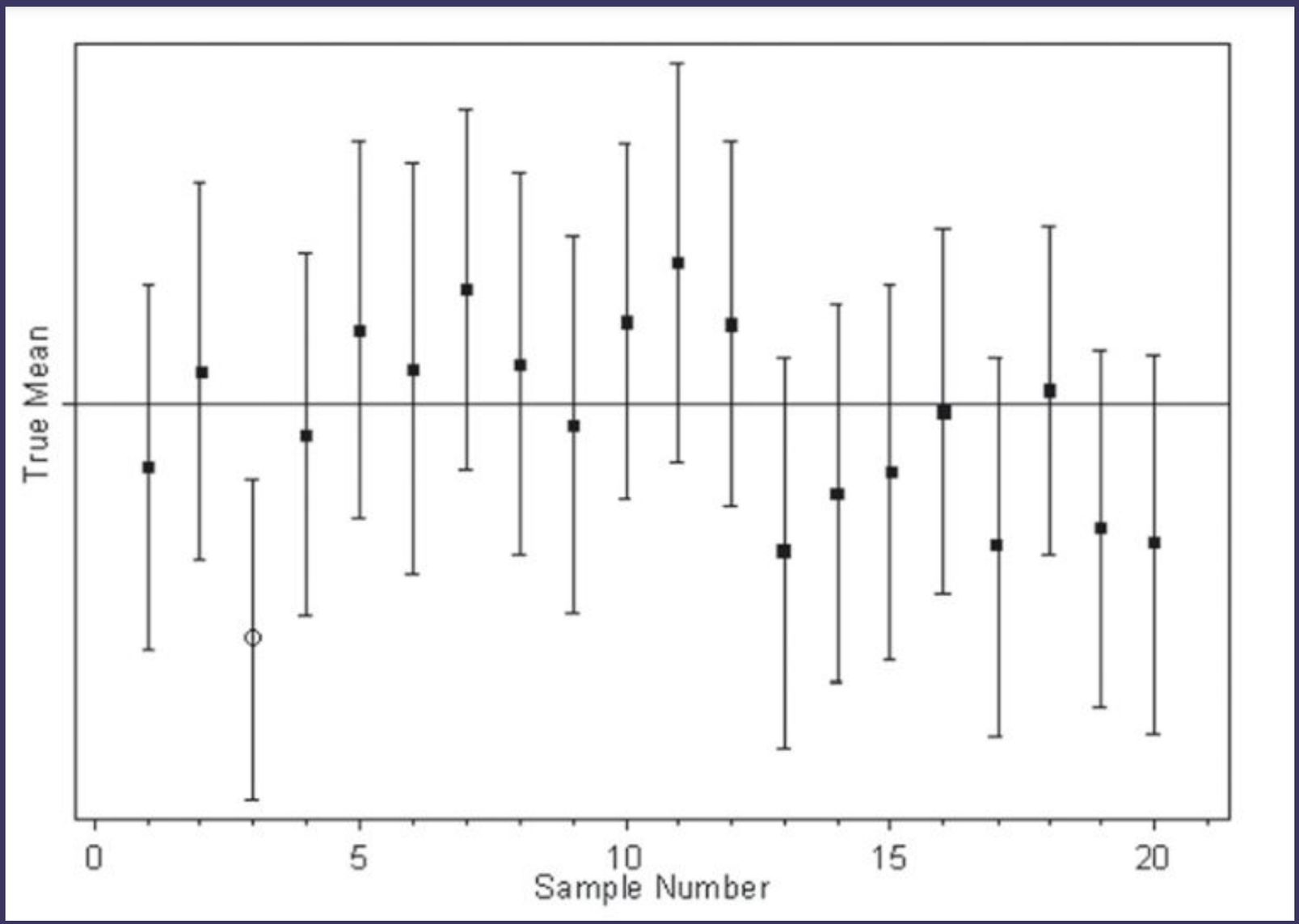
- Estimates precision of hazard ratio

Confidence Interval

95% CI means that if a trial is *repeated* using the *same population* it is estimated that 95% of the intervals would contain the *true value* for said outcome in their interval

- Does **NOT** indicate a 95% chance that the true value is included in a single interval

Confidence Interval



Confidence Interval

Wide CIs = less precise estimates
Narrow CIs = more precise estimates

Regarding hazard ratios - if the CI for an outcome includes the value of 1.00 then the difference **CANNOT** be considered statistically significant

Confidence Interval

Example

HR 0.78 (95% CI 0.57-1.02); p<0.05

- HR and p-value suggests lower risk
- However, CI contains the value of 1.00
  - A statistically significant difference **CANNOT** be claimed

Non-Inferiority Trial

Designed to assess if the active treatment is *no worse* than the control treatment by a predetermined margin

Non-Inferiority Trial

The predetermined margin is called the **non-inferiority margin** (aka NI margin)

- Ex. An NI margin of 1.30 means that in order to claim non-inferiority, the upper limit of the hazard ratio CI must **NOT** include the value of 1.30

- If the CI crosses/touches the NI margin then non-inferiority **CANNOT** be claimed

Non-Inferiority Trial

Non-inferiority trials **cannot** be used to claim superiority (without predetermined testing specified within the trial protocol)

- Likewise, superiority trials **cannot** be used to claim non-inferiority

Analysis Populations

Intent to treat (ITT)

The sample of patients that underwent randomization into the trial

Analysis Populations

Modified ITT (mITT)

The sample of patients that underwent randomization into the trial and met one or more qualifying criteria

Analysis Populations

Per protocol (PP)

The sample of patients that successfully completed the trial

Composite Outcomes

A combination of outcomes reported for a single measure of effect

- Ex. Composite of cardiovascular death, myocardial infarction or stroke

Each component should occur at similar rates and have similar clinical significance

- Ex. Composite of death and minor bleeding would **NOT** be appropriate

Relative Risk Reduction

The difference in event rate of the active treatment group relative to the control group

RRR = 1 - (active/control)

Relative Risk Reduction

Commonly reported for treatment effect

- Easily misinterpreted and tends to overestimate treatment effect

Example

Event Rate A = 10%, Event Rate B = 20%

- RRR = 50% (only 10% absolute difference)

Absolute Risk Reduction

The absolute difference in event rates between two treatment groups

ARR = | control - active |

Less commonly reported than RRR

- Equation can also be used to calculate absolute risk *increase* (ARI)

Number Needed to Treat

An estimate of how many patients would need to receive active treatment to prevent 1 outcome compared to the control treatment

NNT = 1 / ARR

- ARR input as decimal value (10% = 0.1)
- Round NNT **up** to nearest whole number

Number Needed to Harm

An estimate of how many patients would need to receive active treatment for 1 adverse outcome to occur compared to control treatment

NNH = 1 / ARI

- ARI input as decimal value (10% = 0.1)
- Round NNH **down** to nearest whole number

Interpreting NNT & NNH

NNT and NNH are estimates used to illustrate the magnitude of treatment effect in terms of patients, instead of percentages

Interpreting NNT & NNH

NNT < NNH indicates a favorable benefit/risk ratio

- However, the clinical significance of each outcome must be considered
- Duration of trial must also be considered

Only calculate NNT/NNH for statistically significant differences

Level of Evidence

The measure of the quality of evidence from a trial

- **Level I** - RCT with power met
- **Level II** - RCT with power NOT met
- **Level III, IV and V** - observational trials with or without a control group

Grade of Recommendation

Used to rate the strength of the *your own* recommendation

**The higher the level of evidence,
the higher the grade of recommendation**

- **Level I** - Grade A
- **Level II** - Grade B
- **Level III, IV and V** - Grade C

These are *subjective* measures

Purpose

To effectively analyze and interpret trial results in order to create an evidence-based recommendation

References

The following resources were used to form this quick biostatistics review:

Malone PM, Park SK, Malone MJ, eds. Drug Information: A Guide for Pharmacists. Sixth edition.
McGraw-Hill Education; 2018.

Bryant PJ, Pace HA. The Pharmacist's Guide to Evidence-Based Medicine for Clinical Decision
Making. American Society of Health-System Pharmacists; 2008.

Tan, S. H., & Tan, S. B. (2010). The correct interpretation of confidence intervals. Proceedings of Singapore Healthcare, 19(3), 276–278. https://doi.org/10.1177/201010581001900316

Please refer to these resources for more thorough and comprehensive information on the subject.

High-Powered Medicine

Landmark Clinical Trial Reviews

"Good doctors use both individual clinical expertise and the best available external evidence, and neither alone is enough." -*David Sackett, MD*

www.highpoweredmedicine.com