



# What Non-statisticians Need to Know about Statistics in Clinical Trials

Erika Menius, MS

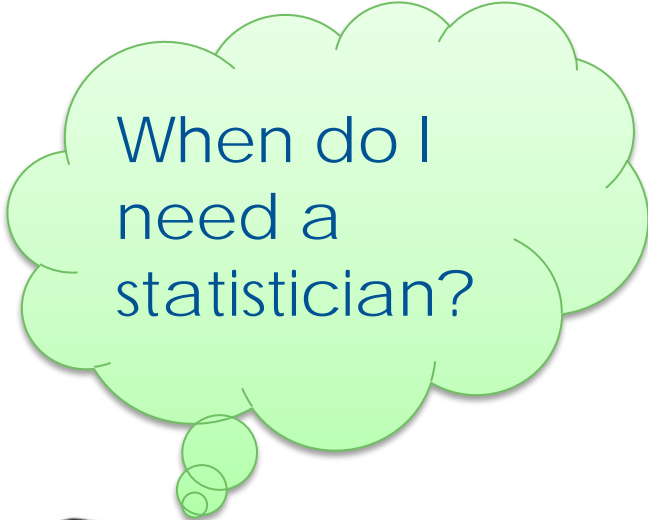
Senior Biostatistician, Rho

# Abstract

Not everyone wants to be a statistician; however, in the clinical trials business, your understanding of statistics can make or break your program. Through real-world examples, webinar participants will learn strategies for choosing appropriate outcome measures, methods for analysis and randomization, and sample sizes. Participants will also learn why planning for data collection is so important as well as tips for collecting the right data to answer your scientific questions. These tips and strategies are what every non-statistician should know about statistics in clinical trials.

# Agenda

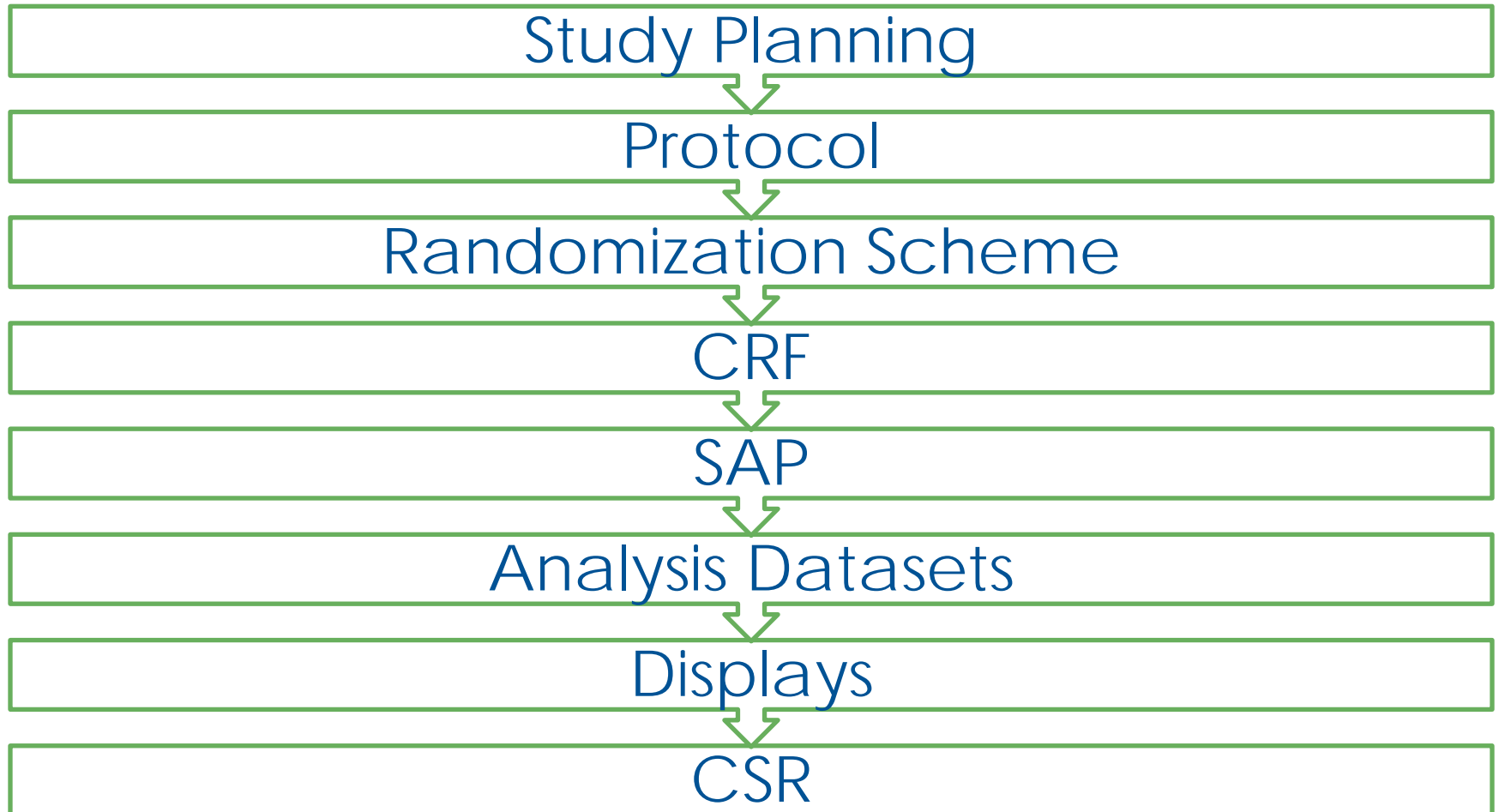
- ✓ Clinical Trial Study Flow
- ✓ Planning Your Trial
- ✓ Sample Size and Power
- ✓ Data Capture
- ✓ Randomization
- ✓ Statistical Analysis Plan
- ✓ Interim Analyses
- ✓ Database Lock
- ✓ Final Analyses
- ✓ CSR



When do I  
need a  
statistician?



# Clinical Trial Study Flow



# Planning Your Trial

What is our goal?

What data do we collect?

How do we test them?

# Planning Your Trial - Example

## Indication:

- OA of the knee

## Goal:

- Show our product is better than placebo

## Data to collect:

- Pain by VAS on 50 foot walk test, multiple collection times

# Statistical Review- Example

Population:

- OA of the knee unilateral, bilateral, age?

Type of data:

- Continuous

Number of time points:

- Repeated Measures

Test:

- LS Means Difference based on Repeated Measures Population Average Model

Sensitivity:

- Unilateral vs. Bilateral, Missing Data

# Planning Your Trial – Blinding/Masking



<http://www.dcscience.net/?p=239>

# Planning Your Trial – Blinding/Masking

## Single Blinding

- The participant doesn't know to which intervention they have been assigned.

## Double Blinding

- The participant and the investigator don't know to which intervention the participant has been assigned.

## Triple Blinding

- The participant, investigator, and monitoring committee do not know to which intervention the participant is assigned.

# Planning Your Trial – Blinding/Masking

## Advantages

- Decrease bias
- Participant response not influenced by knowledge of treatment
- [DB] Investigator preconception does not matter

## Disadvantages

- Patient consent
- Another layer of complexity
- [TB] Patient safety
- Can the study really be blinded?

# Study Populations

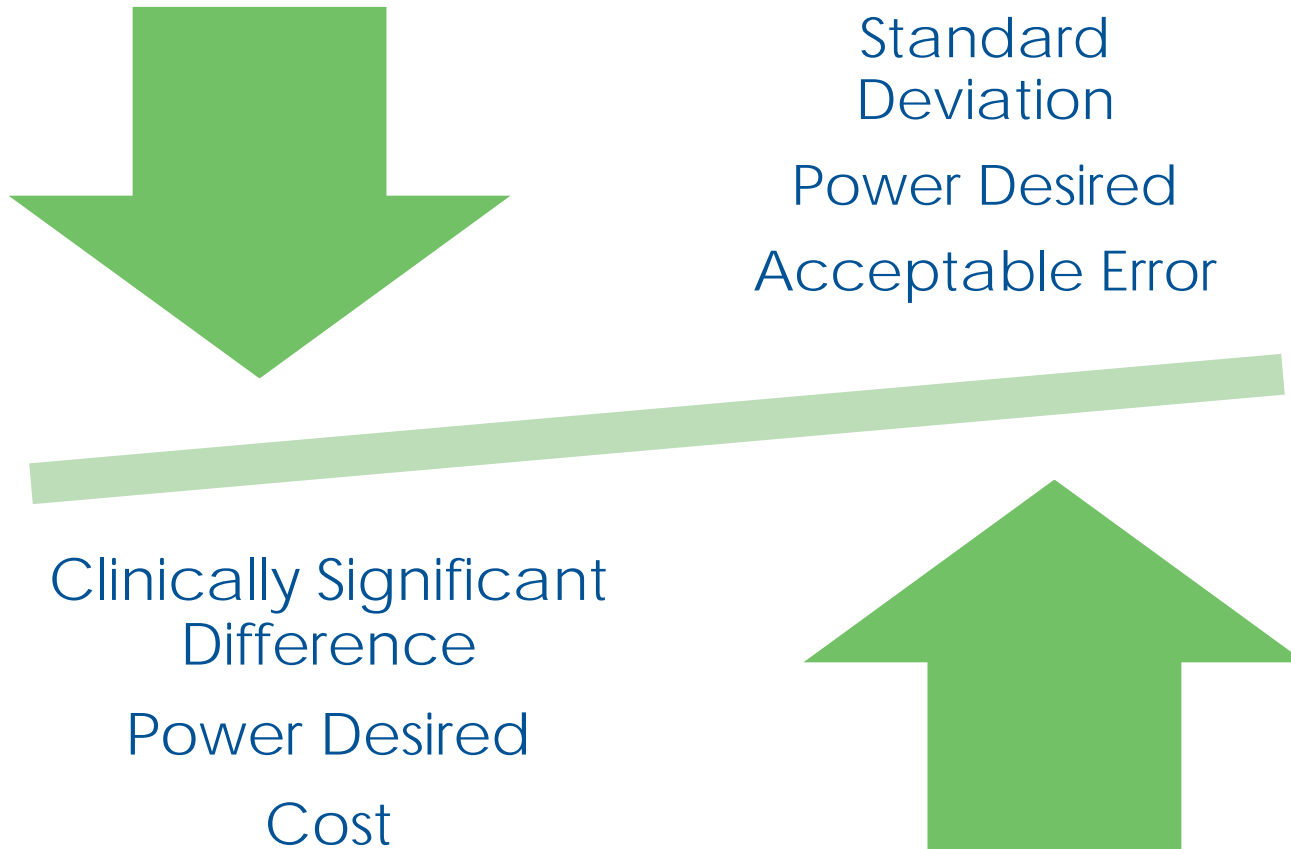
Who do you want in your study?

- Inclusion & Exclusion criteria

Ensure that statistical inference can be made to targeted market population

- Safety Analysis Set
- Full analysis sets (ITT Population)
- Per Protocol
- Depending on draft guidance, Clinically evaluable

# Sample Size and Power



# Sample Size and Power

$$n_1 = n_2 = \frac{2(z_{1-\alpha/2} + z_{1-\beta})^2 \sigma^2}{(\mu_2 - \mu_1)^2}$$

## Decoder Ring

$\alpha$  = Allowable Error (usually 5%)

$1-\beta$  = Power (usually 80%-90%)

$\sigma^2$  = Standard Deviation

$\mu_2 - \mu_1$  = Clinically Significant Difference



# Sample Size and Power

**REMEMBER:**

The lower the allowable error, the bigger the sample size

# Sample Size and Power

**REMEMBER:**

The higher the power, the bigger the sample size

# Sample Size and Power

**REMEMBER:**

The bigger the standard deviation, the bigger the sample size

# Sample Size and Power

**REMEMBER:**

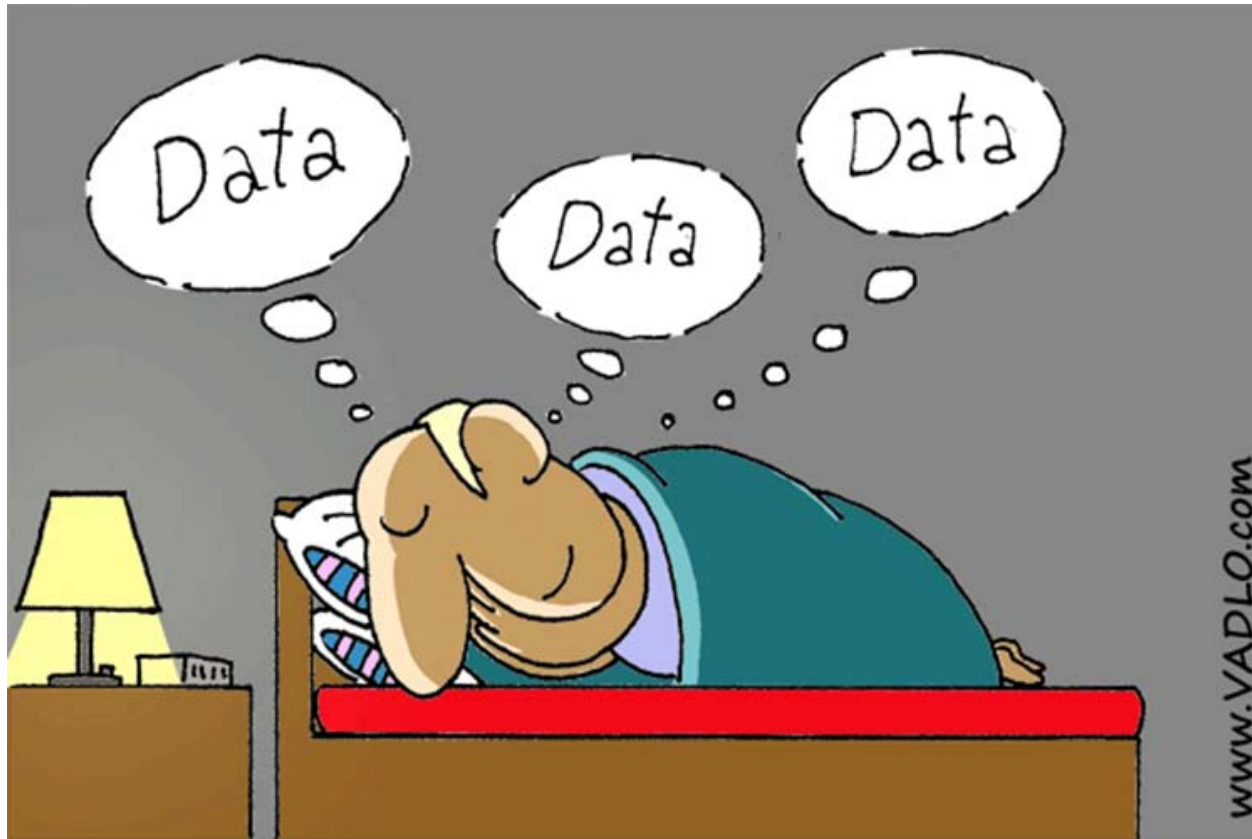
The bigger the clinically significant difference, the smaller the sample size

# Sample Size and Power

**REMEMBER:**

All differences can be “statistically significant” if you have enough subjects, power only for your clinically significant difference!

# Data Capture



The Not-So-Secret life of a PI

# Data Capture

CRF design is integral to capturing the data you need for a successful analysis.

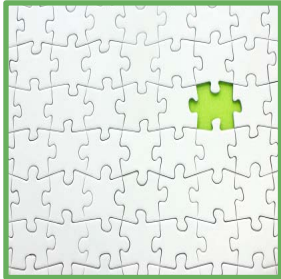
Statisticians need to participate in CRF design to make sure assessments align with analyses!

It's VERY difficult to go back and obtain data after the fact!

Will this study be part of a submission?

- CDASH

# Data Capture – Missing Data

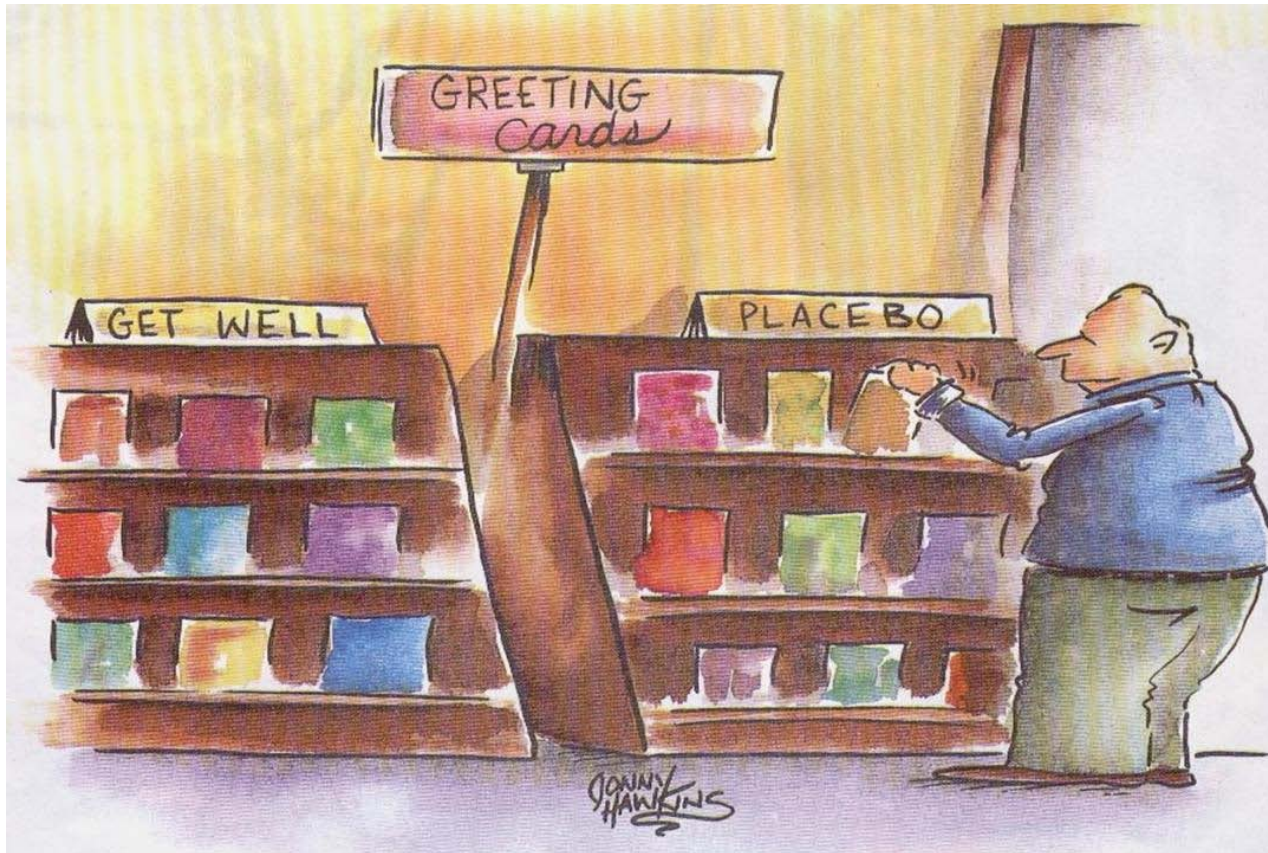


- Potential source of bias
- Minimize through protocol design
- Consult guidance, literature, and your statistician for candidate methods for analysis
- Define and justify the proposed method
- Communicate with client and the internal team

See:

O'Neill, R and Temple, R. "The Prevention and Treatment of Missing Data in Clinical Trials: An FDA Perspective on the Importance of Dealing With It." Clin Pharmacol Ther. 2012 Mar, 91(3); 550-4.

# Randomization



<http://cancergrace.org/cancer-101/2007/02/07/oncology-study-design-and-stats/>

# Randomization



## Reasons

- Reduction of bias
- Sound statistical basis for evaluation
- Produces treatment groups in which the distributions of prognostic factors, known and unknown, are similar

# Randomization - Types

## Simple Randomization

Like flipping a coin

Pro: Easy!

Con: You could randomize everyone to the same group



# Randomization - Types

## Permuted Block Randomization

Randomized by block

Block	1	2	3	4	5	6	7	8	9
Treatments	ABC	CBA	CAB	BCA	ACB	ACB	ABC	CAB	BCA

Pro: Balance across intervention arms

Con: If you know the block size (and it's small), you may be able to guess the next treatment

# Randomization - Types

## Stratified Randomization

If a key factor may affect how an intervention works in a particular group, stratify by that factor.

Can combine this method with permuted block for balance:

Permuted block stratified by baseline pain:

Moderate pain:

**AABB** ABAB **BBAA**



3 blocks  
of size 4

Severe pain:

**ABAB** BBAA **BABA**



3 blocks  
of size 4

# Statistical Analysis Plan

- What data will we use?
- Which participants will be included?
- Exactly how will we analyze?
- Factors affecting analysis?

**Gets down to the nuts and bolts of the statistical analyses**



# Statistical Analysis Plans

- Descriptive Statistics
- T – test and Non-parametric Test (Wilcoxon Test)
- ANOVA and ANCOVA
- Linear Regression
- Linear Mixed Models
- Pattern Mixture Models

Continuous Outcomes



- Pattern Mixture Models
- Missing Data Imputation Methods
  - LOCF
  - BOCF
  - Hot Deck
  - Multiple Imputation

Categorical Outcomes



# Statistical Analysis Plans

- Descriptive Statistics
- T – test and Non-parametric Test (Wilcoxon Test)
- ANOVA and ANCOVA
- Linear Regression
- Linear Mixed Models
- Pattern Mixture Models

Continuous  
Outcomes



- Descriptive Statistics
- $\chi^2$  / Fisher's Exact Test
- CMH test, Odds Ratio, Relative Risk
- McNemar's, Agreement (Kappa)
- Logistic Regression
- Poisson Models

Categorical  
Outcomes



# Statistical Analysis Plans

- Kaplan Meier
- Log Rank Test
- Survival Rates
- Poisson Models
- Cox Proportional Hazard Models

## Survival Analysis

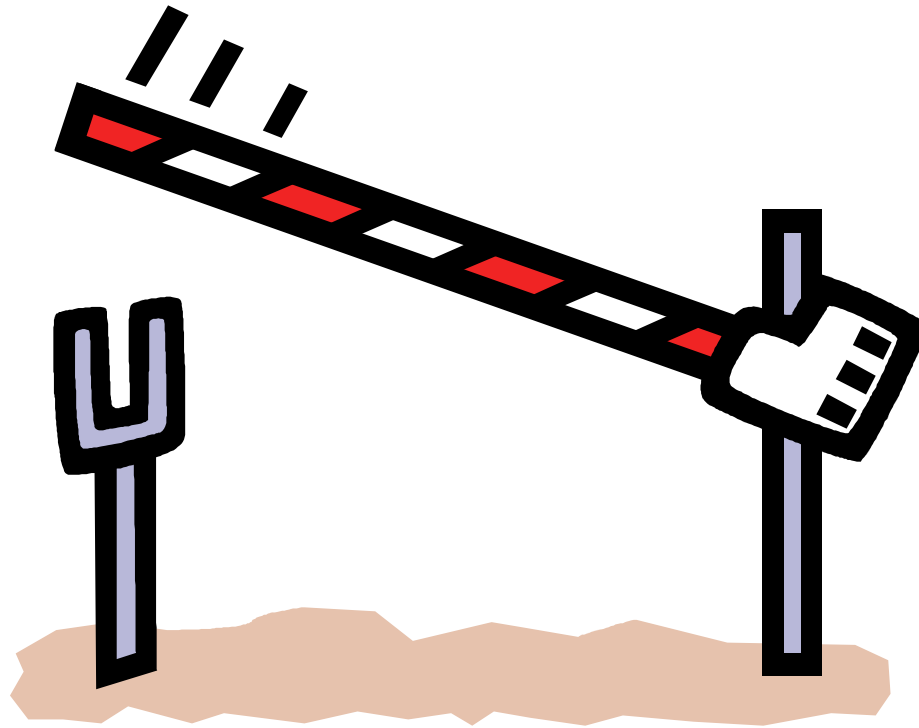


- Pattern Mixture Models
- Missing Data Imputation Methods
  - LOCF
  - BOCF
  - Hot Deck
  - Multiple Imputation

## Sensitivity Analyses



# Interim Analyses



# Interim Analyses

- Is an interim analysis planned?
- What is the purpose of the interim analysis?
- Interim analysis timing and frequency
- Is an unblinded interim team needed?
- What is the data cleaning process for the interim analysis?
- How does this affect  $\alpha$ ?

# Interim Analyses - IDMC/DSMB

Purpose

- Safety

Timing and Frequency

- Based on outcome and safety concerns

Unmasked Team

- Possible, not always necessary

Data Cleaning

- Interim database locks, snapshots

Affected  $\alpha$  ?

- No efficacy data

# Interim Analyses – Sample Size Recalculation

Purpose

- Ensure necessary sample size based on SD assumptions

Timing and Frequency

- Usually just once

Unmasked Team

- Not necessary

Data Cleaning

- Interim database locks

Affected  $\alpha$ ?

- Not if performed in a pooled SD adjustment

# Interim Analyses – Stopping Rules

Purpose

- Efficacy

Timing and Frequency

- Based on primary outcome

Unmasked Team

- Required

Data Cleaning

- Interim database locks

Affected  $\alpha$ ?

- Yes, if study continues

# Interim Analyses – Adaptive Designs

Example: Pruning

Purpose

- Efficacy

Timing and Frequency

- Based on primary outcome

Unmasked Team

- Usually, based on arms involved

Data Cleaning

- Interim database locks

Affected  $\alpha$ ?

- Yes, but not usually well controlled studies (Phase I or II)

# Database Lock and Unmasking



# Database Lock and Unmasking



- All analysis plans should be complete
- Per-Protocol population selection
- Statistician sign off
- Data quality
- Missing data?
- Unmasking

# Final Analyses

- Submission? SDTM and ADaM
- Hypothesis Testing
- Primary/Secondary Outcomes
- Safety Reporting
- Missing Data?



# Clinical Study Report

- Statistical Reporting
- Primary Endpoint Discussion
- What does it all mean??



# Summary

- Determining trial objectives and corresponding endpoints, primary and secondary, is important initial step.
- The type of trial should be aligned with sponsor's clinical plan.
- Determining the sample size early is very important to the projected cost for running the trial.
- Statistical parts of the protocol serve as starting point for all remaining activities.
- Emphasis on design and statistical principles protects the study from bias by specifying the analysis methods a priori.

# References

- ICH-E3: Structure and Content of CSRs
- ICH-E6: Good Clinical Practice: Consolidated Guidance
- ICH-E9: Statistical Principles for Clinical Trials
- ICH-E10: Choice of Control Group and Related Issues in Clinical Trials
- FDA Guidance for Industry: Various Indications
- National Academy of Science Missing Data Guidance

# Contact Us

More webinars available at  
[www.rhoworld.com/rho/resource-center/webinars](http://www.rhoworld.com/rho/resource-center/webinars)

[www.rhoworld.com](http://www.rhoworld.com)

[info@rhoworld.com](mailto:info@rhoworld.com)

919-408-8000