

# Development of Analysis Plans

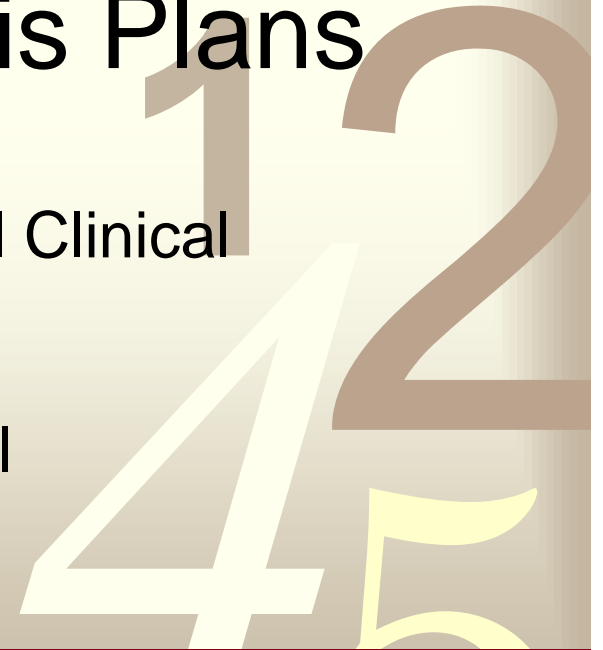
0011

The Fundamentals of International Clinical  
Research Workshop

Mario Chen, PhD.

Family Health International

DMID/ICSSC



# Outline

0011

- Why? When? What?
- Sections of Analysis Plans
  - I. Study Objectives
  - II. Study Design
  - III. Analysis Populations
  - IV. Missing Data
  - V. Participant Follow-up or Disposition
  - VI. Participant Characteristics
  - VII. Primary and Secondary Objectives
  - VIII. Other Planned Analysis
- Concluding Remarks



# Why?

0011

- Helps in the validity and credibility of the results
- There are typically numerous statistical methods to evaluate an outcome, and infinite numbers of analysis sets, covariates, sub-group analyses
- Decisions regarding which methods to use should be pre-specified to avoid potential bias, actual or perceived

# When?

0011

- Pre-specified analysis plans are essential for confirmatory clinical trials. Highly recommended for all other studies
  - Before study initiation
  - Before un-blinding
  - Before initiation of data analysis
- Changes to the analysis plan after study initiation should be clearly documented (v1.0, v2.0, etc.), and should be based on blind reviews of the data
  - Document changes from the protocol summary analysis plan

# What?

0011

- Detailed description of all data analysis steps
- Document all hypotheses, analysis strategy, and assumptions
- Detailed description is essential for confirmatory analysis. Description can be less specific for exploratory analysis
- Analysis plans can be included in the protocol or as a stand-alone document
  - Summary of statistical methods for evaluating primary and key secondary outcomes should be included in the protocol, regardless

# Sections of Plan

---

## I. Primary and Secondary Study Objectives

*Primary: “to assess the effectiveness of a combined TB treatment regimen compared to single drug control regimen”*

*Secondary: “to assess the acceptability of the new dosing regimen”*

- Make clear distinctions between *objectives* (e.g. effectiveness) vs. *outcomes* (e.g. time to cure)

# Sections of Plan

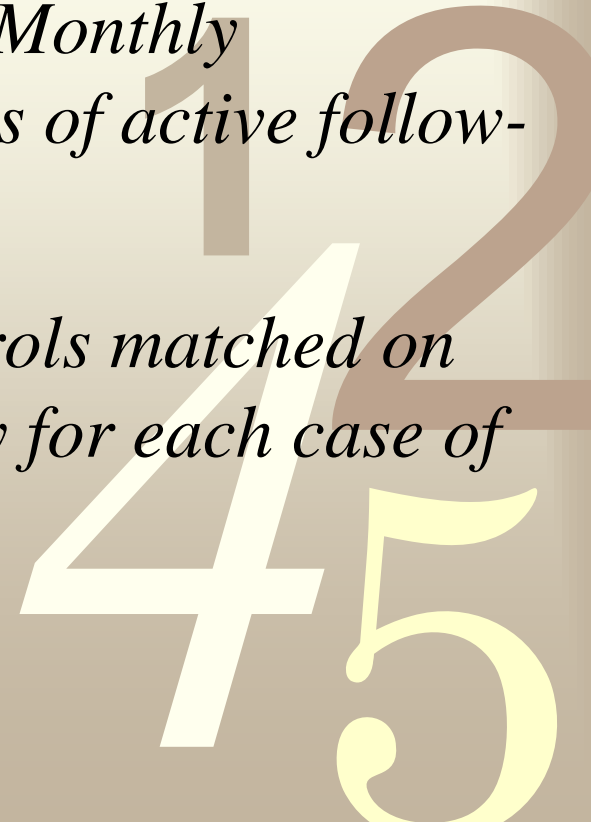
---

## II. Study Design Summary

0011

*“Single-center, parallel group, double-blind trial randomizing 1000 participants to active or control treatments in a 1-1 allocation ratio. Monthly assessment of outcomes for 12 months of active follow-up”*

*“Case-control study, with three controls matched on age and residence, selected randomly for each case of Cutaneous Leishmaniasis identified”*



# Sections of Plan

---

## III. Analysis Datasets

0011

- Full Analysis Set (e.g., all eligible cases and their matched controls) or Intent-to-Treat (ITT): possibly excluding participants with missing key information
- Per-Protocol: participants in compliance with protocol (e.g., excluded matching errors)
- Other pre-specified analysis set(s) of interest for the analysis of the same objective (e.g., no product interruptions) or for other objectives (e.g., safety, subgroups)

# Sections of Plan

---

## 0011 IV. Missing Data

*“Imputation of missing data is not anticipated and will be avoided if possible. Data checks for missing or inconsistent data on key variables (e.g. exposure, key covariates) will be conducted prior to data analysis and decision rules developed as needed and documented”*

- Plan on how you will present the data

# Sections of Plan

---

## V. Participant Disposition

- Numbers of participants included in key analysis sets, person time of follow-up, visits completed, lost to follow-up, etc. described and compared between study groups, sites, etc.
- What statistical methods will be used to make comparisons

# Sections of Plan

---

## VI. Participant Characteristics (or Baseline Data)

- General rule: summarize information for any analysis set used in the main analyses
- Describe what data will be summarized (demographic data, medical history, etc.)
- Will data be summarized by study group, site, or other categories?

# Participant Characteristics (continued)

---

- How will data be summarized (e.g. mean, median, range, SE, category levels)

*If variables are categorized (age 20-24, 25-30, etc.), the rationale for choice of cut-points should be determined prior to initiating comparative analyses*

- Will comparisons be made between study groups? If so, how? For what purpose?

# Sections of Plan

---

## VII. Primary and Secondary Objectives

- Describe analysis sets to be used
- Define variables that are critical to the analyses. Identify outcomes, exposure, interventions, and control variables. Define hierarchy if multiple outcomes
- Specify each primary and secondary hypothesis test, as well as the significance level used, direction of test, whether confidence intervals will be used, etc.

# Primary and Secondary Objectives

---

- Is there a multiple testing issue? If so, how accounted for?
- How will missing data be handled in each analysis. Will it be ignored? Imputation? Sensitivity analyses?
- How will special features of the data be accounted for? Features such as, matching, clustering, complex sampling.
  - Paired t-tests
  - Survey methods
  - Mixed Models or GEE

# Primary and Secondary Objectives

---

- What covariates will be employed

0011

*Covariates to be included should be pre-specified to avoid bias. Otherwise, exploratory. If exploratory, specify method for variable selection*

- Will the adjusted or unadjusted analyses be considered primary?

*“Both adjusted and unadjusted conditional logistic regression models will be fitted. Unadjusted regression will include only the main exposure variable. The adjusted regression will include additional covariates (risk factors and confounders) as listed in Appendix I. The primary evaluation of the association will be the odds ratio and 95%CI from the final adjusted model”.*

# Primary and Secondary Objectives

---

- 0011 What are the planned subgroup analyses (by site, age, etc.). Descriptive only? Limited
- How robustness of primary results will be evaluated. Is there evidence of bias?

*use different analysis sets (with different exclusion rules) and conduct exploratory analyses*



# Primary and Secondary Objectives

---

- Describe precise statistical methodology

0011

*the less precise your methodology, the more exploratory your analysis becomes*

- Include the way in which assumptions will be validated.  
Alternative approaches

- Test for key assumptions (e.g., proportional hazards, carry over, interactions)
  - Power might be low
  - Alternative: Cox models vs. Logrank tests
- Normality and sample size
  - Use exact methods if rare events or small sample size
- Look for minimal assumption methods
  - Logistic Regression vs. Mantel-Haenzel Chi-square



# Primary and Secondary Objectives

---

- May be helpful to include draft table and listing “shells” as an appendix to the analysis plan
- Consider using graphical presentations

0011



# Sections of Plan

---

## VIII. Other Planned Analyses

- Interim Analyses
- Other Objectives

0011



# Concluding remarks

0011

- A pre-specified analysis plan adds credibility to the conclusions of the study
- Include as much detail as possible to limit guessing during analysis implementation
- Check for consistency among the objectives, the analysis plan, and the protocol
- Think about the study report and manuscript
- Obtain feedback

