CONSUMER BEHAVIOR STUDIES STATISTICAL ISSUES

Ralph B. D'Agostino, Sr.

Boston University
Harvard Clinical Research Institute
September, 2006

AIMS OF PRESENTATION

- Identify issues that have major impact on statistical analysis and/or interpretation
- Discuss how to deal with the issues
- WARNING: I think all issues are ultimately statistical issues

TYPES OF STUDIES UNDER DISCUSSION

- Label Comprehension Study (LCS)
  - Label relates to consumer correctly understanding the indication, directions for use and safety warnings
- Self Selection study (SSS)
  - Consumer can appropriately self select or de-select
- Actual Use Study (AUS)
  - Evaluate if consumer can appropriately select or de-select to use drug and then can appropriately use drug while avoiding or addressing safety problems

STUDIES (Examples)

- OTC STATINS
  - Primary: MD consultation within first 2 months
  - Secondary: MD follow-up, LDL-C reduction, Safety, Appropriate lifestyle

- PRILOSEC: 5 Studies, Objective to characterize usage patterns/dosing compliance in OTC setting

OUTLINE

- Study Objectives and Endpoints
- Major Study Design Threats to Validity
  - Populations, setting, other study biases
- Statistical Analysis Issues
  - Data sets, missing data, procedures, sample size, multiplicity, subgroups
- Other Issues and Concerns

STUDY OBJECTIVES/ENDPOINTS

- OBJECTIVES
  - USUALLY **MANY**/OFTEN BEHAVIORAL
- STUDY ENDPOINTS
  - TO CORRESPOND TO OBJECTIVES
  - RESPONSES TO QUESTIONNAIRES OR
  - BEHAVIOR RELATED TO DRUG SELECTION AND USE
- SORT IN GROUPS & PRIMARY,SECONDARY

STUDY OBJECTIVES/ENDPOINTS EXAMPLE (AUS)

- USAGE PATTERN
- SUBJECT UNDERSTANDING and COMPLIANCE

- HEALTH PROFESSIONAL CONTACT
- MONITORING USER'S CONDITION

- SAFETY (DRUG USE and INTERACTIONS)
- EFFICACY

STUDY ENDPOINTS
How to Consider in Analysis

- SEPARATE endpoint evaluation analyzes each variable separately (multiple primary outcomes)
- COMPOSITE has one primary outcome (e.g., percent correct on all components) and then evaluates separate items for consistency as secondary analysis
- Always State Expectation (e.g, 85% success)

MAJOR STUDY DESIGN THREATS TO VALIDITY

- Target Population vs. Study Population
- Setting of the study
- Other Biases

IF THESE ITEMS ARE NOT HANDLED CORRECTLY STATISTICS WILL NOT HELP

TARGET and SAMPLE POPULATIONS

- TARGET POPULATION is the population of potential users
- SAMPLE POPULATION is the population from which study subjects will be drawn
- Each exclusion and inclusion criteria can take us from the target population and limit ability to generalize (My favorites: include people who will cooperate, keep them coming for interviews)

SETTING OF STUDY

- Setting of the study should be as natural as possible and correspond to settings where drug will be used
- Conduct of the study should be as natural as possible (e.g., avoid artificial restrictions in the study such as asking subject to recall label without being able to look at it)

OTHER BIASES

- Study should avoid biases that "lead" subject responses and actions (e.g., asking a question where "desirable" answer is obvious)
- REMEMBER: Most of these studies are non-randomized and also open label. RANDOMIZATION AND BLINDING not available to reduce or eliminate biases

STATISTICAL ANALYSIS ISSUES

- Analysis Data Sets
- Drop Outs/Missing Data
- Statistical Procedures
- Sample Size
- Multiplicity
- Subsets

STATISTICAL ANALYSIS ISSUES

- Often we do not have a control or comparison group (one arm study)

- Analysis is focused on measured behavior and presented as a percent (e.g., percent comprehending label correctly) for the single arm of the study

ANALYSIS DATA SETS

- ANALYSIS DATA SET PRESPECIFIED BEFORE ANALYSIS OR BLIND BROKEN

- INTENTION TO TREAT (ITT)
- PRE PROTOCOL (valid cases)
- ADEQUATE USER
  - TO STATISTICIAN: THE HOPELESS CASE

DROP OUTS/MISSING DATA

- NEED TO MINIMIZE DURING STUDY
- NEED TO UNDERSTAND MECHANISM (completely at random, at random, non-ignorable)

- MISSING DATA SHOULD BE IMPUTED?
  - LAST OBSERVATION CARRIED FORWARD?
  - IMPUTATION DEPENDS ON MECHANISM

- **SENSITIVITY ANALYSIS ESSENTIAL**

STATISTICAL ANALYSIS PROCEDURE #1

CONFIDENCE INTERVAL

- All outcome variables should have a pre-stated level of expectation (e.g., we expect 85% success)
- .95 CONFIDENCE INTERVALS USED to present data

POINT ESTIMATE ± MARGIN OF ERROR

.87 ± .03

Is expected in the confidence interval?

STATISTICAL ANALYSIS PROCEDURE #1

- If the above is employed SAMPLE SIZE is easy to compute

$$N = (1.96/2 \times \text{Margin of Error})^2$$

N for .03 margin of error is 1067

N for .05 margin of error is 385

MORE PRECISE FORMULAS EXIST

STATISTICAL ANALYSIS PROCEDURE #2

HYPOTHESIS TESTS

- State a level of outcome considered a failure

  A value that we want to "rule out"

- Example: H0: $p \leq .80$ vs. H1: $p > .80$ at level of significance 0.025

STATISTICAL ANALYSIS PROCEDURE #2

- Sample Size formulas and software for obtaining sample size exist that require

1. Value to rule out 0.80
2. Value expected 0.85
3. Need about 500 subjects (for power 0.80)

4. For 0.80 and 0.90 N is about 100

SAMPLE SIZE FORMULA

$$N = ((A + B)/C)^2$$

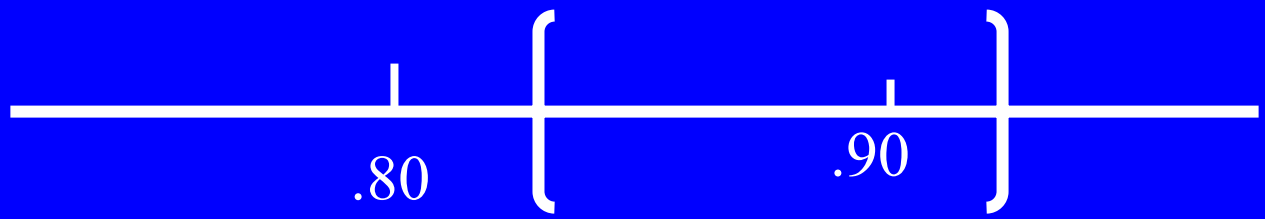
$A=1.645\text{Sqrt}(p_0(1-p_0))$, p_0 rule out value

$B = 0.84\text{Sqrt}(p_1(1-p_1))$, p_1 expected outcome

$C = p_0 - p_1$

STATISTICAL ANALYSIS PROCEDURE #3

- Above two procedures can be combined where a 0.95 confidence is computed and it is desired to have the value to rule out below the lower confidence limit (also desire to have expected value in the interval)



STATISTICAL ANALYSIS ISSUES: MULTIPLICITY

- Multiple primary endpoints require statistical adjustment procedures (e.g., confidence level (coefficients), alpha spending, multiple looks at data) (SAMPLE SIZE increase)
- Adjustments needed within set of endpoints for given objective and across sets of objectives
- Adjustments for multiple objectives may be negotiable (e.g., Compliance and safety)
- Adjustments decided a priori

STATISTICAL ANALYSIS ISSUES: MULTIPLICITY

- One useful way to control multiplicity effect is to deal with composite variables as primary variables. Look at components as secondary analyses

STATISTICAL ANALYSIS ISSUES: SUBGROUPS

- Secondary analyses should include examination of subgroups for consistency

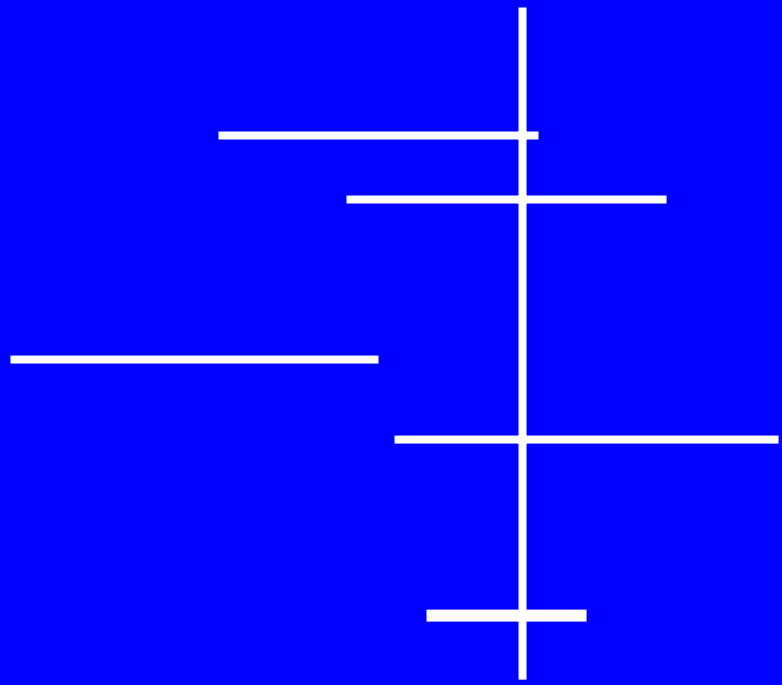
> 65yrs

≤ 65

Literacy low

Literacy high

OVERALL



COMPARISON or CONTROL GROUPS

- CONTROL GROUPS OFTEN NOT INVOLVED
- COMPARISON GROUPS NOT INVOLVED

- These would supply internal validity for study

- Could test variations of labels and questionnaires

- Analysis would be extended to comparison of study arms

ISSUES (DESIGN/CONDUCT/ANALYSIS)

- **STUDY OBJECTIVES**
- **TARGET POPULATION**
- **SAMPLE POPULATION**
- **ENDPOINTS**
- SELF REPORT vs. VERIFIABLE DATA
- **CONTROL GROUPS**
- STUDY DESIGN (BIAS)
- STUDY DESIGN (SAMPLES)
- LENGTH OF STUDY
- **DATA ANALYSIS**
- TRIAL MONITORING
- **DROP OUTS**
- **MISSING DATA**
- SAFETY
- **SUBSETS**
- CLINICAL SIGNIFICANCE

CLOSING COMMENTS

- MANY STATISTICAL ISSUES, NONE UNIQUE TO CONSUMER BEHAVIOR

- Thing out objectives
- Have “good” sample, avoid biases
- Decide carefully on outcomes
- State expectations, size study accordingly