# General Statistical Challenges for CNS Devices: A Regulatory Perspective

Gregory Campbell, Ph.D.
Director, Division of Biostatistics
Center for Devices and Radiological Health
Food and Drug Administration
greg.campbell@fda.hhs.gov

International Society for CNS Trial Methodology, Washington, DC Feb. 26, 2008


# Outline

- Introduction to Trial Planning in Design, Conduct and Analysis
- Bias and the Selection of the Control
- Masking as a Way to Decrease Bias
- Decreasing Variability
- Issues in Trial Conduct
- Problems in the Analysis and Interpretation

# Introduction

- High standard of regulatory research
  - Precisely specified primary endpoint and Statistical Analysis Plan
  - Sample size to assure adequate power
  - Inspection of PMA data
- The nature of devices is different than drugs—evolution not discovery.
- Different regulatory standard for devices

# Endpoints

- Patient Reported Outcomes (PRO), Physician Assessed Outcomes or Third-Party Assessed Outcomes can serve as primary endpoints.
- For PRO:
  - http://www.fda.gov/cder/guidance/5460dft.htm
  - VAS for Patient Reported Outcomes
- For ordinal scales, each person may be operating on his/her own scale (and may be different from time to time). Means across patients may not be informative.
- Change from baseline but baseline may be highly variable
- The value of secondary endpoints

# Randomized Controlled Trial

- Gold standard of clinical trials
- Control is either placebo (sham), nocebo or an active control

# Sham (or Placebo) Control

- The treatment arm may improve over baseline but so may the control.
- In such a case it is the difference in the effectiveness of the two arms that measures the true effect of the device.
- The sham arm may improve due to the placebo effect.
- Placebo effect may be different for devices (Kaptchuk et al, *J Clin Epi* 2000)

# Placebo Effect and the Choice of Control

- Sham controls for devices are sometimes more difficult to arrange than placebos in drug trials.
- NIH Conference on the Placebo in 2000
- Placebo effect
  - Hawthorne effect
  - Expectation of benefit
  - Regression to the Mean

# Regression to the Mean (RTTM)

- Galton: children's heights of tall parents regress toward the mean
- Rookie effect in baseball is an example.
- Danger of choosing extreme patients since they may regress to the mean.

# Examples of the Placebo Effect in Sham Surgeries

- Very relevant for many CNS studies involving pain or function (or depression).
- NIH Parkinson's trial of 40 patients using fetal tissue transplantation in an RCT with a sham surgery as the control using a patient reported outcome

  Freed et al *NEJM* 2001;**344**:710-719.
- Arthroscopic surgery for knee OA

  Moseley et al NEJM 2002;347:81-88.

# Nocebo Control

- Control receives no additional therapy (and they know it) other than the usual Standard of Care (SOC)
- No expectation of benefit, so if there is a difference between the new therapy and the nocebo, is it due to the new treatment or the expectation of benefit?
- This control introduces a bias.

# Active Control: Non-Inferiority Studies

- Uniquely regulatory for use in active control trials
- Effort is to show that new therapy (or diagnostic) is no more than slightly worse than another medical procedure (and since the comparator is effective then so is the new therapy)
- Null hypothesis: the new device is inferior to the standard by at least some small amount delta.
- Alternative: the new device is not inferior by more than delta (it may actually be superior)
- Non-inferiority studies can reduce the bias (in the selection of the control) but can be very difficult to perform.

# The Importance of Masking

- Can the patients feel or sense the therapy? (Can they guess it?)  Does Informed Consent suggest they may experience some effect as a result of the therapy?
- Is there a questionnaire to ask if the patients are aware of (or can guess) their treatment?
- If it is impossible to mask the treatment, this introduces a bias.
- If there is a sham control but patients can sense the treatment when present, it is really a nocebo trial.

# One-Arm Studies: Patient as Own Control

- Two approaches: before-after or on-off
- Before-after can be plagued with bias issues in term of all 3 types: Hawthorne, expectation of benefit and in some cases RTTM as in:
  - select extreme patients and sometimes they improve regardless of treatment
- For on-off studies the order can be randomized (as in a crossover design, worrying about the washout) but there can still be regression to the mean if the treatment is successfully masked. Another similar design is high-low instead of on-off.

# One-arm Studies: Historical Control

- The additional challenge of using historical control data
  - Change in medical practice or in the population
  - Historical control needs to be identified and agreed upon prospectively
  - Need for patient-level data for propensity score analysis; this control can pose risks to the sponsor.
- The use of historical controls can introduce bias due to the placebo effect in the one-arm study (Hawthorne, expectation of benefit, RTTM)

# Large Variability

- When the observations are highly variable it is very difficult to detect the signal, to see the effect if it is there.
- If variability is due to known covariates it may be possible to reduce the variance through covariate modeling but it is difficult to plan studies with this as the primary analysis (could stratify)
- Center-to-center variability
- Learning curve for the surgeon
- Different skill levels for the surgeons (if an implant)

# Decreasing Variance (Increasing Precision)

- Stabilize the baseline, either by studying stable disease or by multiple baseline observations.
- Increase the sample size to detect a reasonable clinically significant difference.
- Measure change from baseline and compare between two groups, the treatment and the control.
- More homogenous population
- Number of failed treatments

# Conduct of the Trial

- Dropouts
- Crossovers of patients from one arm to the other
- Rescue meds
- Challenges in how to analyze such data

# The Bane of Missing Data

- Missing data can destroy a great design, reducing a wonderful RCT to an observational study.
- Good design is to plan to follow up vigorously and when that fails to have a strong plan of how to treat the missing data in the analysis
- Types of missingness
  - Crossover, death (suicide), fail to appear for one evaluation time or fail to appear in the correct window, rescue meds, drop out of study altogether
- Intention-to-treat principle
  - Superiority versus non-inferiority
- Example: Cool-Cap

# The Current Rage of Adaptive Design

- What are adaptive designs?
  - Sequential monitoring to stop early for safety or effectiveness
  - Sample size re-estimation
  - Drop an arm
  - Predictive modeling (Bayesian)

# The Use of Bayesian Statistics

- Role of CDRH
- Two approaches
  - Use of prior information
  - Build predictive models
- Bayesian trials tend to be adaptive.
- Draft guidance: http://www.fda.gov/cdrh/osb/guidance/1601.html

# Diagnostic Studies

- Most diagnostic products are regulated in the US by CDRH. This includes all kinds of imaging systems as well as *in vitro* diagnostics and *in vivo* systems too.
- Statistical Guidance on Reporting the Results from Studies Evaluating Diagnostic Tests:

  http://www.fda.gov/cdrh/osb/guidance/1620.html
- Intention-to-diagnose

# Good Reporting Practice for Nonpharmacologic Treatments

- CONSORT Statement for Nonpharmacologic Treatments (2008)
- *Ann Int Med* 2008;**148**:295-309

# Neurological Devices Panel

- FDA Advisory Committee
- Transparency and unconflicted advice
- Information concerning Neurological Devices Panel on the FDA website for medical devices http://www.fda.gov/cdrh/panel/

# The Challenge of Many CNS Device Studies

- CNS device studies are difficult to design, conduct, analyze and interpret.
- There may be no perfect design.
- Good planning is absolutely crucial.
- CDRH encourages device sponsors to come in to FDA early and discuss plans at IDE or pre-IDE meetings.


Center for Devices and Radiological Health
CDRH