Image Interpretations Variability and Bias

Jyoti Zalkikar, Ph.D.
Office of Biostatistics
CDER/FDA

Image Interpretations "Reads"

- Integral part of those clinical trials in which Medical Imaging is used for
  - Diagnosis
  - Measurement of response to therapy
- Performed by Experts known as "Readers"
- Inherently Subjective
- Lead to endpoints that are subject to Bias and Variability

Two Main Sources of Variability

- Intra- Reader
  - Variation due to case difficulty
- Inter-reader
  - Variation due to Reader skill

- Differential Image Acquisition, reader training and information available to readers all contribute to both these sources of variability

Reducing Variability

- Standardize Image Acquisition
  - Imaging charters
  - Machines and Phantoms
- Standardize Reader Training
  - Blinded Independent Central Read
- Assure uniformity of Information available to readers
  - Blinded Independent Central Read

Blinded Independent Central Read

- Reduces Bias by enabling control over the information available to readers
- Reduces Bias by enabling implementation of randomization in the “read design”
- Likely to produce high quality data due to controlled, transparent setting.
- Enables quantitative measurement of Intra- and Inter-reader variability

Intra- Reader Variability
Reader performance Index

Statistical Measures

- Kappa Statistics
  - For categorical data
  - Commonly used and Well-developed (can use weights)
  - Pre-specified evaluation in most imaging charters
- Intra-class Correlation
  - For continuous data
  - Commonly used and Well-developed
  - Assumes linear relationship
- Concordance Correlation
  - For continuous data
  - Corrects for scale and shift differences in readers

Inter-reader Variability

- Statistical Measures : Same as Intra-reader variability
- Expected in Diagnostic Imaging
  - Its absence
    - may cast a doubt on independence of readers
    - May limit the ability to generalize results
- Not well-understood in Therapeutic trials with Imaging Endpoints
  - Its presence leads to
    - Analytical difficulties
    - Reservations about the efficacy of the new treatment

Manage Reader Discordance

- Use 3 independent blinded central readers and majority read
  - Pre-specify the algorithm to generate majority read in the protocol/Analysis Plan
- Analyze the data for each reader separately and show treatment success for each reader
- Use 2 reader-adjudicator paradigm

Pre-specify one method as primary and use others for sensitivity analyses

Mitigate using pre-specified ROI (tumors, vessels, regions) and rigorous reader training (measuring, scoring)

Site versus Central Read

- Two Schools of thought:
  - Blinded read is an unnecessary expense, site reads should be used for primary efficacy analysis
  - Uncontrolled site reads with associated confounding bias cannot substitute for blinded reads
- Blinded Independent Central Reads - A norm in diagnostic imaging

Site Versus Central read

- A Central blinded adjudication committee of experts who interprets all data, clinical as well as imaging, and provides consensus “Read” (the endpoint result)
  - Works for Incidence rate (Anti-coagulation trials)
  - Precludes investigator bias
  - Lessens measurement error
  - May not work for some time to event endpoints such as PFS (Informative censoring ?)

Bottom-Line

- Prospective Planning (sample size)
- Prospective Data Collection
- Prospective Analysis plan

Show that data provides robust (unbiased) evidence of treatment success after accounting for various sources of variability