From Symptomatic Effect to Disease Modifying Effect: *A statistical methodology overview in Alzheimer Disease trials*

Richard Y. Zhang, Ph.D.

Director, Biometrics
Pfizer Inc

Statistical Challenge in AD

The following factors should be considered in designing and analyzing the AD trials

- Long term and progressive disease
- Multiple clinical symptoms require multiple endpoints (patients, physicians and caregivers)
- Statistical significance vs. clinical relevance of treatment effect in short-term trial
- Long term trial may have greater clinical relevance but issues of drop out and ethics of placebo treatment

Study Design I

- 2 randomized, double-blind, placebo-controlled parallel trials.
- **Treatment duration**: 3-6 month
- **Efficacy** : 2 primary endpoints
  - Cognition (ADAS-cog) AND
  - Either Global (CIBIC-plus) OR Function (ADCS-ADL)

Leber P. *Guideline for the Clinical Evaluation of Antidementia Drugs*. Division of Neuropharmacological Drug Products. FDA 1990

Statistics Methods I

- Sample Size:150-200/group (based on effect size of less sensitive scale, usually the Global scale)
- Continuous Variables: Mean change from baseline (ANCOVA)
- Categorical Variable: CMH test
- Analysis Population: ITT-LOCF

Cognition: ADAS-cog

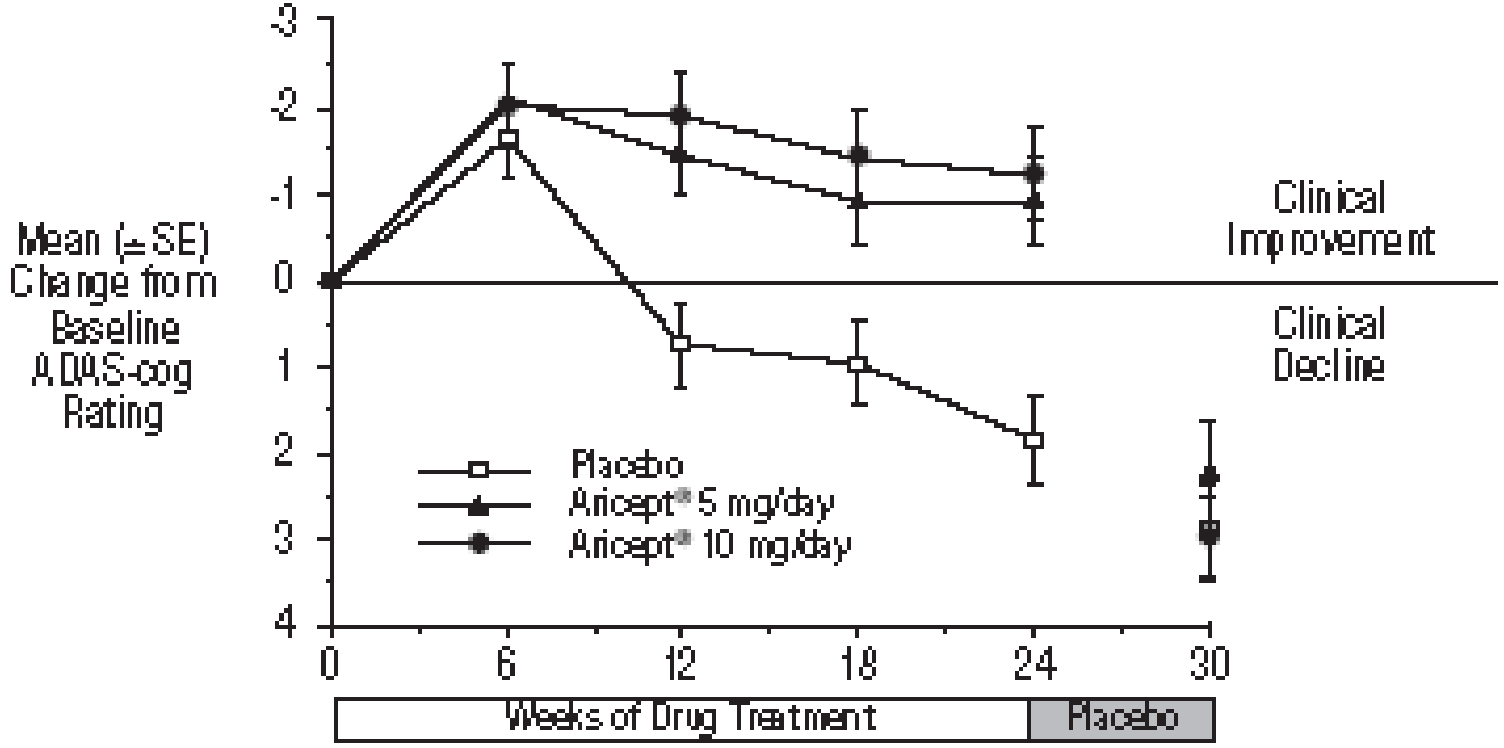


Figure 1. Time-course of the Change from Baseline in ADAS-cog Score for Patients Completing 24 Weeks of Treatment.

Donepezil PI, Eisai/Pfizer.

Zhang R, 2008 ISCTM Autumn Meeting

Global: CIBIC plus

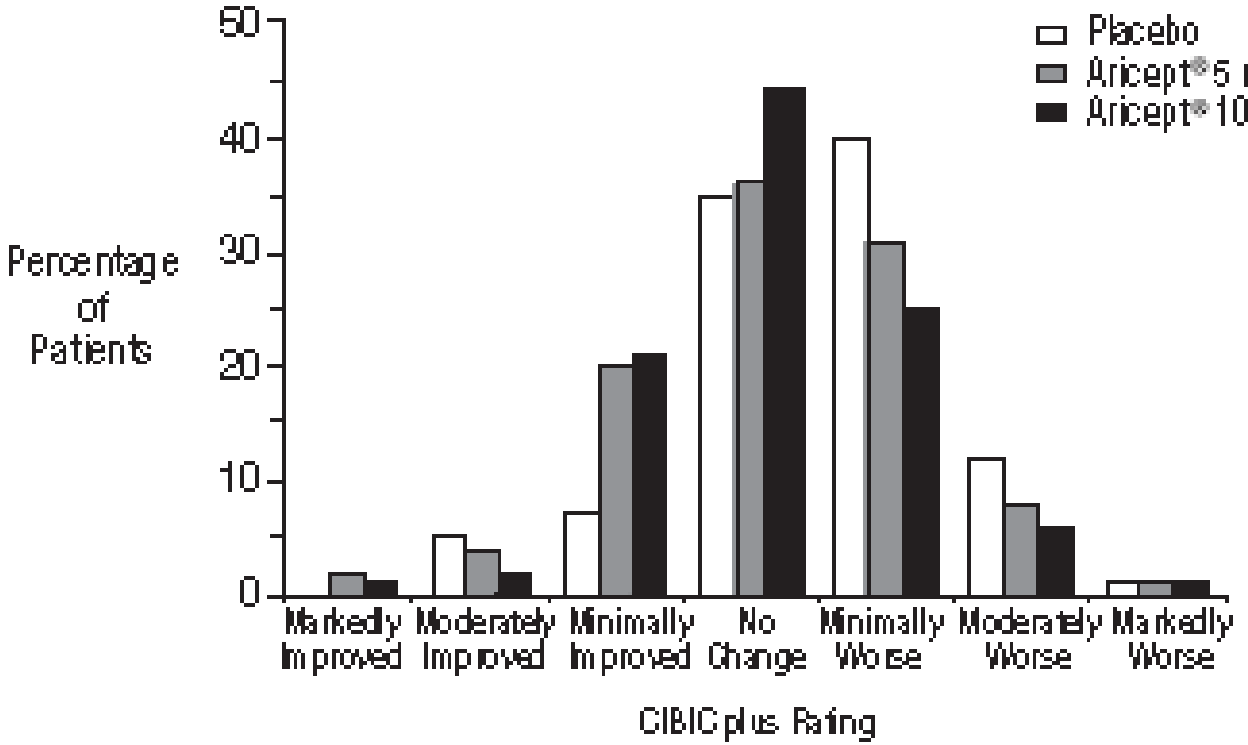


Figure 3. Frequency Distribution of CIBIC plus Scores at Week 24

Donepezil PI, Eisai/Pfizer.

Zhang R, 2008 ISCTM Autumn Meeting

Summary I

- Trials are not designed to demonstrate an effect altering the course of the underlying dementia process.
- Current approved drugs to treat AD are indicated to provide only symptomatic benefit.
- Newer compounds are being developed that considered to be potentially capable inducing beneficial structure changes on the underlying progression of AD. New approaches has to be explored.

Completed/Ongoing Studies of Disease-Modifying Agents in AD

Compound	Phase III Studies
Bapineuzumab (Elan)	18 month DB parallel design, 800 (2-arm) and 1250(4-arm), Primary endpoint: ADAS-cog and DAD ApoE genotyping, MRI
LY450139 (Lilly)	21 month DB parallel design 1500 (3-arm) with Randomized start mechanism*, MRI, PET Primary endpoint: Cognition and Function
Flurizan (Myriad)	18-month DB parallel design, 12-month IA, 1800(2-arm) Primary endpoint: cognition and function
Lipitor (Pfizer)	18 month DB parallel design 641 (2-arm) Primary endpoint: Cognition and Global (Lead) MRI
Alzhemed (Neurochem)	18-month DB parallel design 1052 total (3-arm) Primary endpoint: ADAS-cog and CDR-SB, MRI

Zhang R, 2008 ISCTM Autumn Meeting

Study Design II

Difference:

- Treatment duration: at least 18 month
- Sample size: 350-800/arm
- Placebo group: most subjects are on background AD therapies
- Biomarker endpoints: Added

Similarity:

- Co-primary endpoints
- Traditional parallel design

Statistics Methods II

Difference:

- Mixed Effect Model vs. ANCOVA:
  - –Slope analysis in additional to point estimate
  - –Missing data algorithm other than LOCF

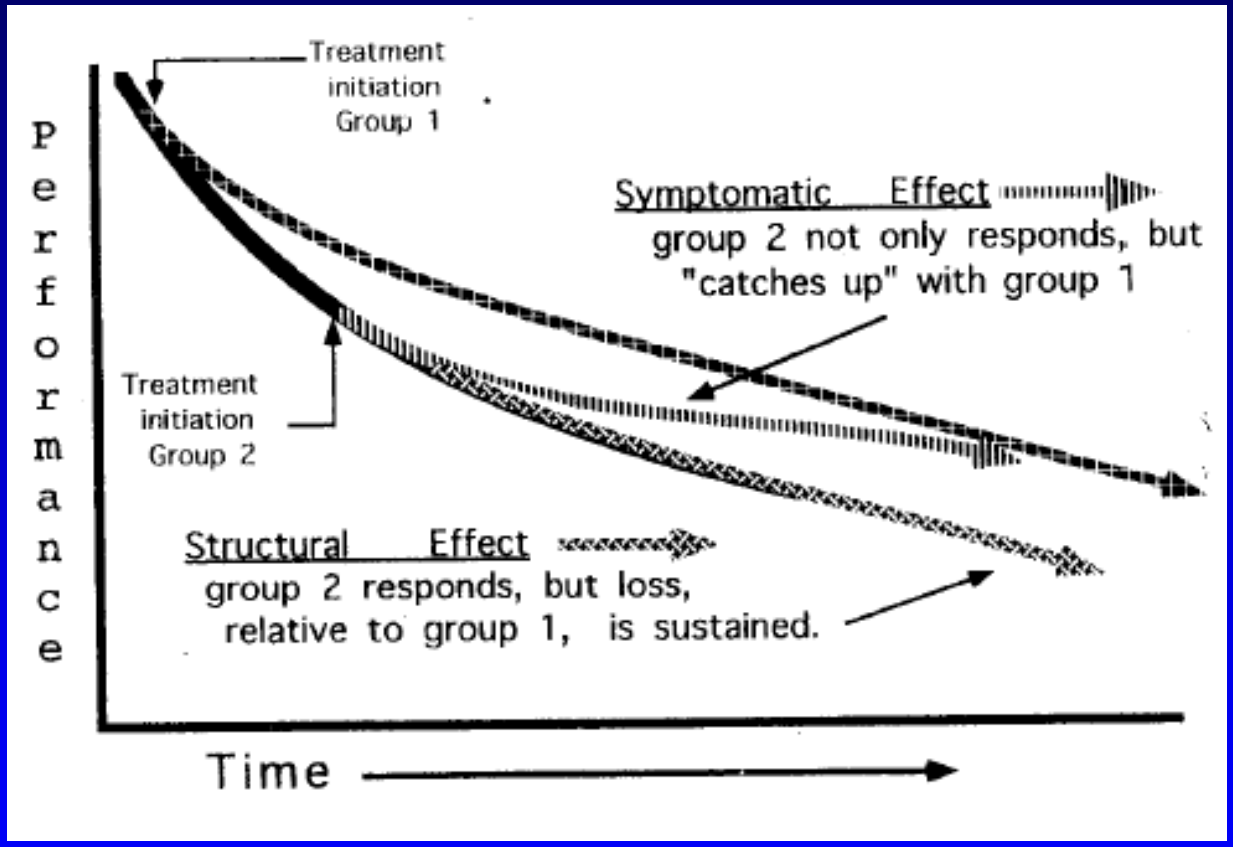
Similarity:

- Primary Analysis: a single primary time point analysis at study endpoint

Summary II

- There is no break-through on clinical trial design or statistical analysis in newer trials vs. old trials.
- The statistical methodologies of disease modification are much underdeveloped and need more attention from statistical communities.

Randomized Start Design



Under the null hypothesis of the randomized start design, the investigational drug has no effect whatsoever on the neuropathologic process. Accordingly, a delay in its administration should not have any lasting deleterious effect on patients denied immediate access to its benefits

Leber, ADAD 1997

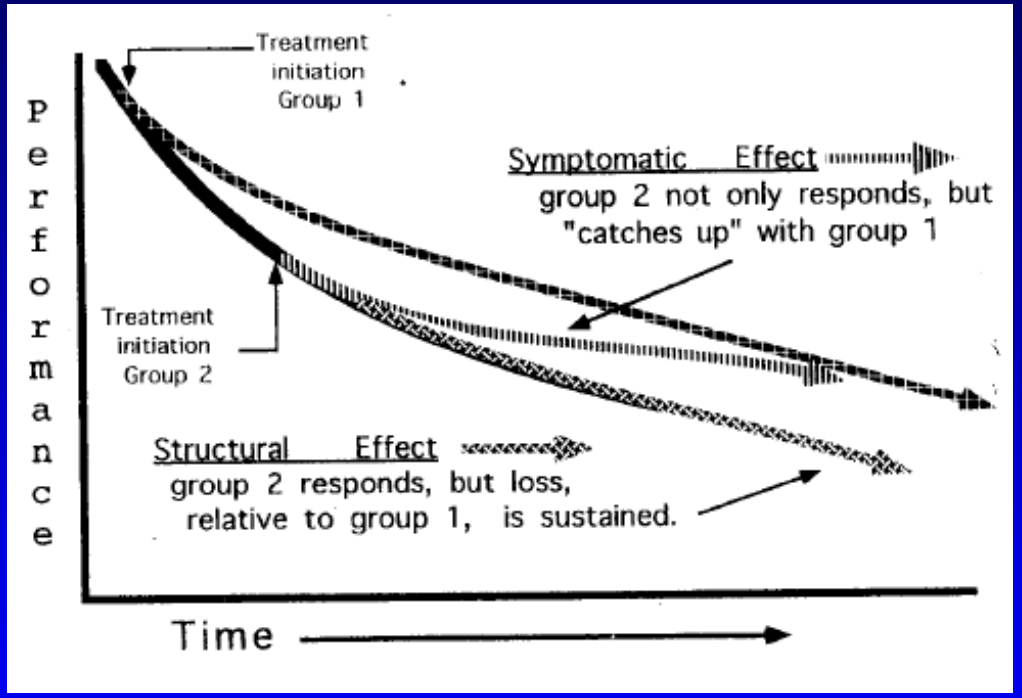
Different Views

- Katz (2004) *"If well designed or conducted, they (randomized withdrawal or start design trials) can be interpreted unambiguously as demonstrating a structure effect. Unfortunately, these trials are rarely performed."*
- Mani (2004) *"There are a number of unresolved matters in regard to both study designs. These include sample size (which is assumed to have to be large), methods of comparing treatment group, duration of observation, and frequency of assessment. To my knowledge, there is no instance where either of these designs has been successfully applied to in a clinical trial of a putative disease-modifying drug in AD"*
- Cummings (2006) *"Given the complex issues surrounding clinical trial design, implementation, and analysis it is unlikely that data derived from clinical trials alone will be sufficient to conclusively demonstrate disease modification."*
- Mohs (2006) *"There is no study design that can unambiguously determine that a drug has an effect to slow the progression of AD, but some design features that would at least consistent with an effect. "*

Methodological issues

Statistical Inference:

- What to use estimate disease progression (a point/points, a slope, trajectory or AUC?)
- What is appropriate statistical test?
- What is minimum frequency of interval measurements?
- Duration of the trial?
- What is proportion of placebo phase vs. delay-start phase?
- How to calculate sample size?

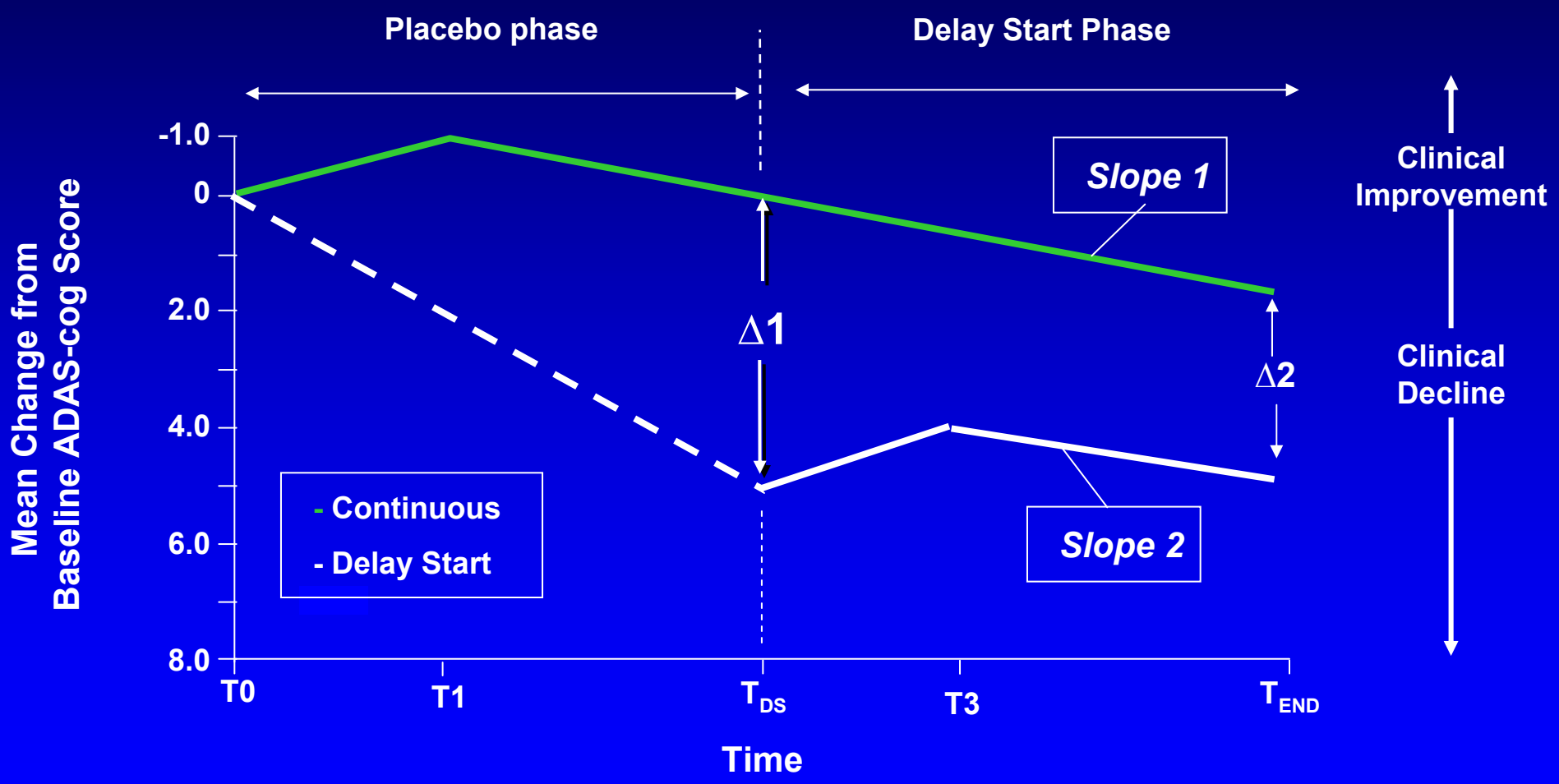


Adjustment

- Comparability: The disease progression is not linear over severity
- Drop-out

Zhang R, 2008 ISCTM Autumn Meeting

Statistical Conception of Randomized Start Design using ADAS-cog as Endpoint



Zhang R, 2008 ISCTM Autumn Meeting

Disease Modifying Criteria

- Criteria 1: At the end of placebo phase, there is a treatment effect observed
- Criteria 2: At the end of delay start phase, there is also a treatment effect observed
- Criteria 3: Further at the end of delay start phase, such treatment effect should be sustained

- Statistics Test 1: $\Delta 1 > 0$
- Statistics Test 2: $\Delta 2 > 0$
- Statistics Test 3: *Slope 1 = Slope 2*
  (i.e. Non-inferiority test:
  Slope 1-Slope 2 > -ε)

How to construct a slope and ε?

Construct Slope and ε

- A linear slope could be constructed between time T_3 and time T_{end} using more than 2 observation points
- Given Δ2 is DM effect, ε is set to 15% (1/7) of annualized Δ2. In other word, the DM treatment effect will still be maintained at least 50% level after 3-4 years delay start treatment
- The Linear Mixed Effect Model could be used to test the slope difference by *Treatment by Time* Effect, where *Time* is continuous. If the lower bound of one-sided 90% Confidence Interval > - ε, the claim could be established

Conceptual Study Design

- Treatment Duration:
  - The placebo phase should be longer enough so that no more symptomatic effect will be added to the continuous treatment arm
  - The delay start phase should be at least equal length of placebo phase so the symptomatic effect from delay start group will be fully worn out
  - *12-month placebo, 12 month delay start, total 24 month*.
- Assessment Time Interval:
  - Assess more frequently in a short time interval close to the end of delay start phase
  - *Assessment at month 21, 22, 23,24*

Sample Size

- Based on modeling of prior cognitive trial data

	Effect Size	Sample Size (2-arm, total)
$\Delta 1$	0.4	200
$\Delta 2$	0.2	800
ε	0.1	2000*

*one-sided α=0.1

Summary

- It is practical that Randomized Start Design could be implemented in AD trials with proper statistical tests constructed.

Questions/Comments

- Proposed DM criteria?
- Methodology of the slope?
- Selection of ε ?
- Setting of study design?
- Sample Size estimation?