U.S. FOOD & DRUG ADMINISTRATION

Increasing Sample Size in Clinical Trials: Considerations Beyond Power

Andrew Potter, PhD
Division of Biometrics 1, OB/OTS/CDER, FDA
ISCTM 14th Annual Scientific Meeting – Feb. 20, 2018
Washington, D.C.

Disclaimer

This presentation reflects the views of the author and should not be construed to represent FDA's views or policies.

What is a Treatment Effect?

- The effect of the a drug compared to a control – active or placebo
  - In this presentation, consider a short term depression trial with the treatment effect measured by difference in mean HAM-D scores
- $\delta = \mu_{trt} - \mu_{plb}$, where δ is the treatment effect, μ_{trt} is the treatment mean HAM-D score, and μ_{plb} is the placebo mean HAM-D score
  - Estimated by difference in adjusted means

Variability of Treatment Effect

- Variability of treatment effect depends on:
  - Sample size (N)
  - Standard deviation of HAM-D scores $(\sigma_{trt}, \sigma_{plb})$
    - Assumed to be equal $(\sigma = \sigma_{trt} = \sigma_{plb})$
    - Assumed to be homogeneous across all study sites
- Standard error $\mathrm{se}(\hat{\delta}) \sim {}^{\sqrt{2}\,\sigma}/_{\sqrt{n}}$
  - In theory, precision increases with increasing sample size
  - Events in a specific trial can violate assumptions

Potential Considerations relating to Sample Size and Precision

- Sufficient to detect expected treatment effect
  - Bayesian criteria: Probability of Success
  - Frequentist criteria: Power
- Study population should be large enough to provide reasonable robustness of results
- Study population should have enough sample to cover important subgroups
  - Regional and Geographic
  - Gender, Race, etc.
- Sufficient patients exposed in the drug development program to detect a safety signal
  - Discussed in ICH E1

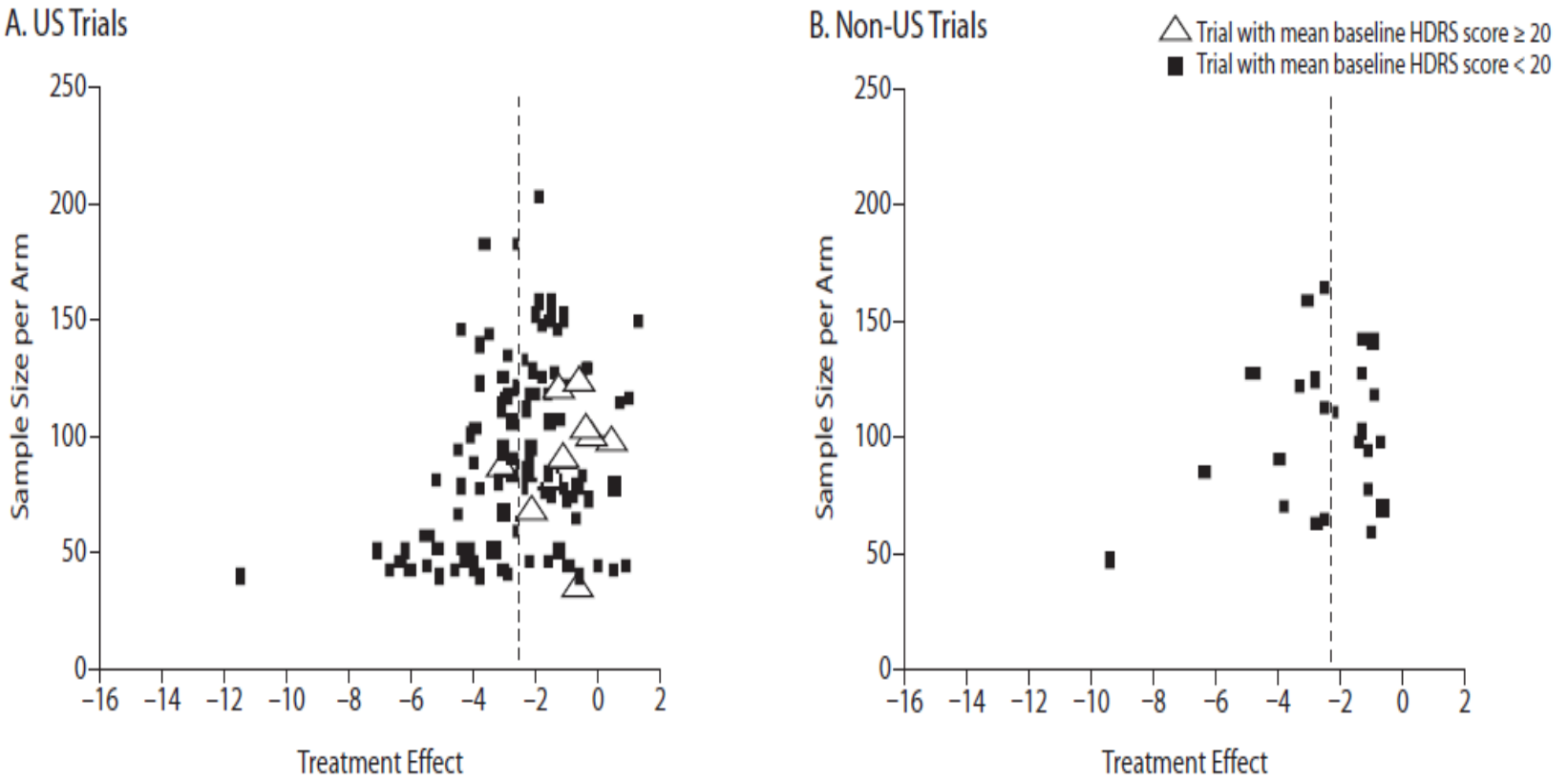
Traditional Power Analysis: How Large a Trial to Detect a Treatment Effect

Standardized Effect Size $d = (\mu_{trt} - \mu_{plb})/\sigma$	Total Sample Size
0.1	4203
0.2	1051
0.3	467
0.4	263
0.5	169
0.6	117
0.7	86
0.8	66

- Total sample size for a trial design to detect treatment effect δ assuming
  - Common σ in both arms
  - 90% power
  - 5% alpha – 2 sided
  - Equal allocation
  - Calculated using EAST 6

Treatment Effect Variation Compared by Sample Size: MDD

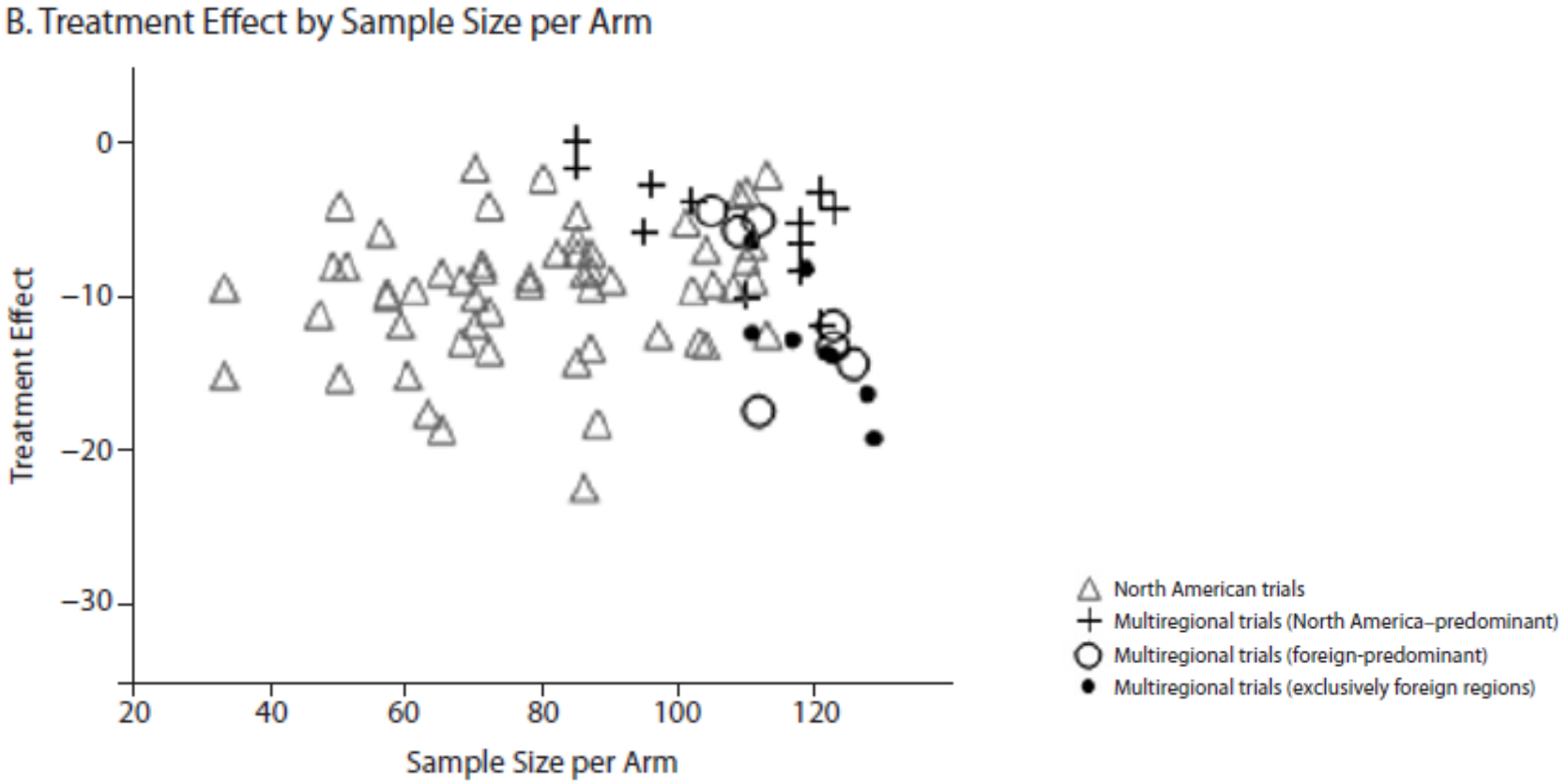
Figure 1. Treatment Effect Relative to Placebo (drug-placebo difference) Based on Mean Change From Baseline to Endpoint (LOCF) in HDRS Total Scores in US and Non-US MDD Trials[a]



[a]Dashed lines indicate mean treatment effect.
Abbreviations: HDRS = Hamilton Depression Rating Scale, LOCF = last observation carried forward, MDD = major depressive disorder.

Source: Khin et. al., *Exploratory Analyses of Efficacy Data From Major Depressive Disorder Trials Submitted to the US Food and Drug Administration in Support of New Drug Applications. J Clin Psychiatry 2011; 72(4): 464-472.*

Treatment Effect Variation Compared by Sample Size: Schizophrenia



Source: Khin et. al., *Exploratory Analyses of Efficacy Data From Schizophrenia Trials in Support of New Drug Applications Submitted to the US Food and Drug Administration. J Clin Psychiatry 2012; 73(6): 856-864.*

Potential Issues of Smaller Trials

- Has an adequate sample of patients been exposed to allow an assessment of safety?
- How robust are the treatment effect estimates?
  - Smaller trials have greater inter-trial variability
  - Another small trial may not show the same treatment effect
- How should the trial results be extrapolated from narrow study population to the general population?
  - Unsure how to label drug for general population
  - May not have sufficient sample size in important subgroups
  - Study population may not capture within region variation

Potential Issues of Larger Trials

- Has trial conduct been consistent across a large number of sites?
  - Larger trials typically have more sites compared to smaller trials
  - Site to site variation in training of study procedures
  - Varying site quality
- Has standard of care changed during trial enrollment?
- How homogeneous is the trial population?
  - Larger trials usually enroll a broader population than smaller trials
  - Leads to larger between patient variation


FDA
U.S. FOOD & DRUG
ADMINISTRATION