An R-Shiny Application for Design of Comparative Clinical Endpoint Bioequivalence Studies

Wansu Park[1,2], Myong-Jin Kim[1], Liang Zhao[1], Lanyan (Lucy) Fang[1]

[1]Division of Quantitative Methods and Modeling, Office of Research and Standards, Office of Generic Drugs, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA

[2]Oak Ridge Institute for Science and Education, Oak Ridge, TN, USA.

♦ Background

- For certain drug products such as complex locally acting drugs, bioequivalence (BE) to a reference listed drug (RLD) may be evaluated via comparative clinical endpoint BE studies.
- The design of a comparative clinical endpoint BE study is generally a blinded, randomized, and parallel study with various statistical test methods (Fig 1).
- A placebo arm is usually included in order to demonstrate that the study is sufficiently sensitive to detect product differences in the patient population enrolled in the study.
- However, these studies could be challenging due to a large sample size, insensitiveness (e.g., small effect size), and cost and time associated with conducting such studies.

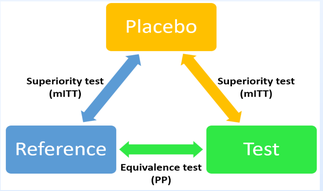


Fig 1. Statistical tests in comparative clinical endpoint bioequivalence studies.
mITT: Modified Intention to Treat; PP: Per Protocol

♦ Objective

- To facilitate the design of a comparative clinical endpoint BE study in terms of
  - Estimate sample size and power
  - Determine study duration, such as the time point to assess BE
  - Evaluate the study sensitivity
  - Streamline the study design to save time and budget

♦ Methods

- An R-Shiny based application (app) was developed to facilitate the design of BE studies with clinical endpoints in terms of sample size and study duration.
- The R and RStudio were used for programming and app testing.
- For continuous endpoints, the *PowerTOST* package was used for a sample size and power determination. For binary endpoints, published statistical equations were used. [1]
- For illustrative purpose, two hypothetical case studies were used.

♦ Results

Case #1 : **Sample** Size/Power Estimation for BE Assessment with Clinical Endpoints at a Prespecified Timepoint

Table 1. Input Values for Case #1

Type of Input	Contents	Value
Efficacy	Proportion of success for RLD	63%
	Proportion of success for Test	63%
	Proportion of success for Placebo	27%
Design features	Enrolled subjects qualified for the mITT	67%
	mITT subjects qualified for the PP	82%
	Desired Power	80%
	Acceptable Type I error rate	5%

Fig 2. Screenshot of Shiny Application for Input Values (left) and Final Sample Size Tab Output (right) in Case #1

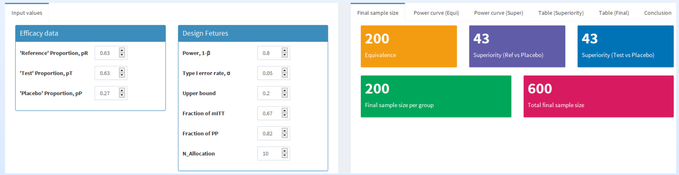
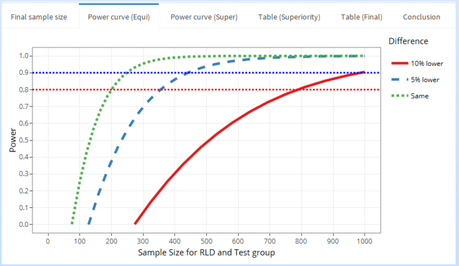
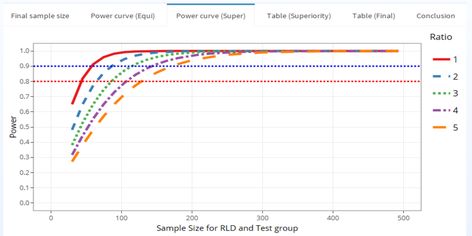


Fig 3. Power Curve of Equivalence Test in Case #1



- Explore the study power based on differences of treatment response rate between the RLD and Test products.
- Assuming the same, 5% and 10% lower response rate for the test product as compared to the RLD, sample size were estimated to be 200, 355, 789 per group for 80% power of study, respectively.
- Based on the assumption of same efficacy, sample size for 80%, 85%, 90% of power were 200, 220, 250 per group, respectively.

Fig 4. Power Curve of Superiority Test in Case #1



Ratio 1: RLD:Test:Placebo=1:1:1; Ratio 2: RLD:Test:Placebo=2:2:1; Ratio 3: RLD:Test:Placebo=3:3:1
Ratio 4: RLD:Test:Placebo=4:4:1; Ratio 5: RLD:Test:Placebo=5:5:1

- Due to the large effect size (RLD 0.63 vs Placebo 0.27), superiority test is easy to pass.
- Imbalanced designs (i.e., smaller placebo arm as compared to active treatment arms) are explored.

Table 2. Summary of Sample Size with Different Allocation Ratio and Power in Case #1

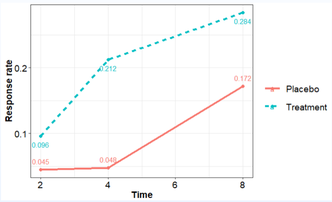
	Allocation	Power	Active	Placebo	Total	Sample_size
1	9	0.8	224	25	473	RLD:Test:Placebo = 224:224:25
2	10	0.8	239	24	502	RLD:Test:Placebo = 239:239:24

	Allocation	Power	Active	Placebo	Total	Sample_size
1	8	0.9	269	34	572	RLD:Test:Placebo = 269:269:34
2	9	0.9	299	34	632	RLD:Test:Placebo = 299:299:34
3	10	0.9	329	33	691	RLD:Test:Placebo = 329:329:33

- Final sample size and allocation ratio were determined to satisfy both the equivalence and superiority test.
- Final sample size recommendation was based on the total sample size to pass BE.
- Final recommended sample size were RLD:Test:Placebo = 224:224:25 and 269:269:34 when power was 0.8 and 0.9, respectively.

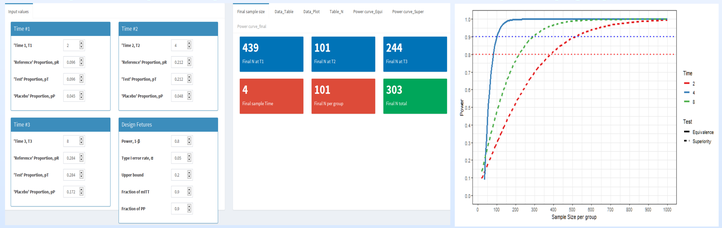
Case #2 : How to Determine the Time Point to Assess BE

Fig 5. Treat Response Rate at Various Timepoints (Days, Case #2)



	time	ref	test	placebo
1	2	0.096	0.096	0.045
2	4	0.212	0.212	0.048
3	8	0.284	0.284	0.172

Fig 6. Screenshot of R-Shiny Application for Input Values (left), Final Sample Size (middle) and Power Curve at Each Timepoint (right)



- Final sample size at each timepoint was determined by combined superiority and equivalence tests.
- Best timepoint and sample size were 101 per group at Day 4 in this case study.

♦ Conclusion

- The developed R-Shiny app could be used an efficient tool to assist comparative clinical endpoint BE study design.

♦ References

[1] Tu D. *J Stat Comput Simul*. 1997;59(3):271-290.

♦ Acknowledgements & Disclaimer

Dr. Park was supported in part by an appointment to the Research Participation Program at the U.S. FDA, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and FDA. The This poster reflects the views of the authors and should not be construed to represent FDA's views or policies