Novel Group Sequential Comparative Clinical Endpoint Bioequivalence Study

Wanjie Sun[1], Zhongying Xu[2]
[1]FDA/CDER/OTS/OB/DBVIII, [2]University of Pittsburg

Acknowledgement: OTS/OB ORISE Program

Abstract

Background:
A comparative clinical endpoint bioequivalence (CCEBE) study is often used to establish bioequivalence (BE) between a locally acting generic drug (T) and reference drug (R), where a typical pharmacokinetic (PK) BE study is uninformative. CCEBE, however, are very expensive to generic drug applicants due to a much larger sample size (compared to PK BE studies) required to achieve a desired power. How to reduce sample size and cut cost for generic drug applicants while not jeopardizing statistical principles is a pressing task for both regulators and applicants. Group sequential design (GSD) is often used in new drugs to allow studies to stop early for efficacy or for futility. However, as of today, there is no published literature providing statistical methods to apply GSD to CCEBE studies.
Purpose
In this work we aim to apply GSD to CCEBE studies to reduce sample size and cut costs for the generic drug applicants.
Methods:
One challenge in applying GSD to CCEBE is that two efficacy tests and one equivalence test are involved in establishing BE, which poses more statistical challenge than an efficacy study for new drugs or PK BE study where only equivalence is involved. In this work, we use Yang and Sun (2019)'s exact method based on a multivariate non-central t distribution to estimate the initial sample size and adopt Pocock (1977) and O'Brien & Fleming(1979)'s method and alpha spending functions to conduct interim analysis at 50%, 75%, and 90% of data completion.
Results:
Our simulation results demonstrate that the proposed novel method controls the Type 1 error rate under the nominal level, attains reasonably high power, and has minimal bias. Most importantly, the proposed method reduces the average total sample size of the study as compared to a fixed study design in general.
Conclusion:
Our work is the first application of GSD to CCEBE studies. The proposed method can help generic drug applicants reduce sample size while attaining reasonably high power without jeopardizing Type 1 error control. This will help applicants cut costs and make CCEBE studies more affordable, hence improving the accessibility of generic drugs to the public.

Introduction

- Comparative clinical endpoint BE studies are very expensive to generic applicants due to a large sample size required to achieve a desired power.
- Adaptive design is often used to optimize study design for new drugs.
- Potvin et al (2018) and other authors proposed several adaptive sequential methods for PK BE studies.
- However, for clinical endpoint BE studies, few work has been done.
- In 2019, Zhu and Sun proposed the first two-stage adaptive clinical endpoint BE studies, i.e., when applicants are not certain about the inter-subject variance of drug products during study design, sample size can be re-estimated at interim based on the estimated variance from the first stage.
- Besides sample size re-estimation, another typical way to reduce sample size is group sequential design which allows applicants to stop the study early if efficacy (for new drugs) is established at the interim analysis, but has not been applied to clinical endpoint BE studies yet.
- In our work we aim to propose a novel group sequential comparative clinical endpoint BE studies such that a good study can be stopped early if BE is established (likewise for futility) at interim, which would reduce sample size and cut cost for the applicants.

Materials and Methods

Hypothesis Tests in Comparative Clinical Endpoint BE Studies:

Superiority tests ($\alpha = 0.025$)*:*

$$H_{01}:\ \mu_T - \mu_P \le 0\ \ vs.\ \ H_{11}: \mu_T - \mu_P > 0 \qquad (1)$$

$$H_{02}:\ \mu_R - \mu_P \le 0\ \ vs.\ \ H_{12}: \mu_R - \mu_P > 0 \qquad (2)$$

Equivalence tests ($\alpha = 0.05$)*: Two One-sided Tests*

$$H_{03}:\ \frac{\mu_T}{\mu_R} \le \theta_1\ \ vs.\ \ H_{13}: \frac{\mu_T}{\mu_R} > \theta_1 \qquad (3)$$

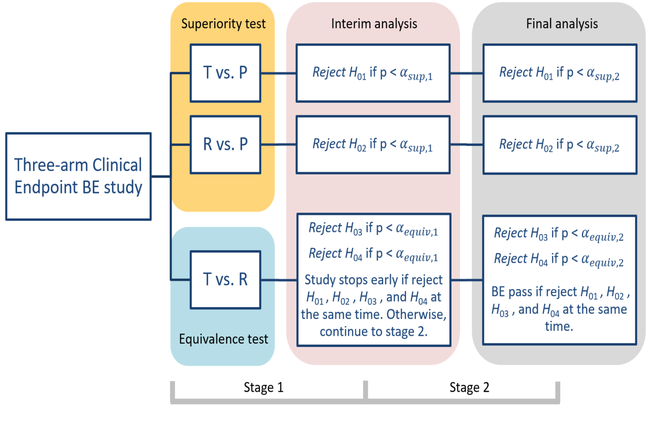
$$H_{04}:\ \frac{\mu_T}{\mu_R} \ge \theta_2\ \ vs.\ \ H_{14}: \frac{\mu_T}{\mu_R} < \theta_2 \qquad (4)$$

- Initial sample size is calculated using Yang and Sun (2019)'s exact method based on a multivariate non-central t distribution.

Table 1. Stopping boundaries for interim and final analyses using Pocock and O'Brien-Fleming methods

Alpha-spending methods	Timing	Superiority test $\alpha_{sup,1}$	Superiority test $\alpha_{sup,2}$	Equivalence test $\alpha_{equiv,1}$	Equivalence test $\alpha_{equiv,2}$
Pocock	50%	0.0147	0.0147	0.0304	0.0304
	75%	0.0167	0.0167	0.0346	0.0346
	90%	0.0191	0.0191	0.0393	0.0393
O'Brien-Fleming	50%	0.0026	0.0240	0.0088	0.0467
	75%	0.0100	0.0219	0.0238	0.0431
	90%	0.0164	0.0214	0.0350	0.0428

Figure 1. Proposed Group Sequential Comparative Clinical Endpoint BE Study Flow Chart



- Stopping for futility can also be added, which does not inflate type 1 error rate.

Results and Discussion

Figure 2 Familywise Type 1 Error Rate by μ_P Stratified by Timing of Interim Analysis, μ_R=0.5, σ=0.3, Assumed geometric mean Ratio (GMR)=0.95,True GMR=0.8.

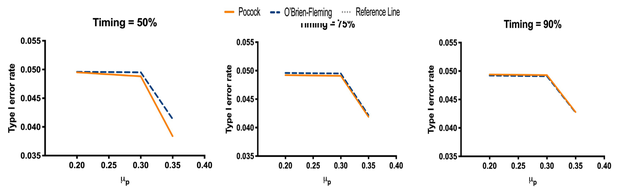


Figure 3 Power by μ_P Stratified by Timing of Interim Analysis μ_R=0.5, σ=0.3, Assumed GMR=True GMR=0.95.

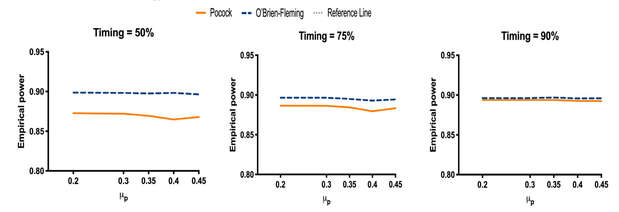
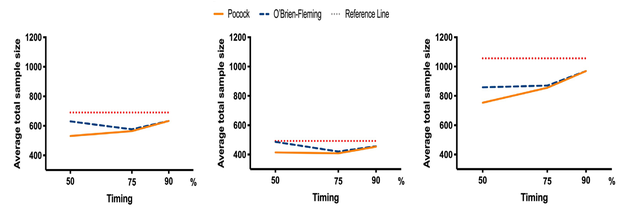


Figure 4 Average Total Sample Size by Timing of Interim Analysis

μ_p=0.35, $\mu_R = 0.5, \sigma$=0.3



μ_P	0.2	0.3	0.35	0.4	0.45
Fixed Design Sample Size	684	684	690	1068	9084

Figure 5 Proportion of Studies Continue to Stage 2 by Timing of Interim Analysis

μ_p=0.35, $\mu_R = 0.5, \sigma$=0.3

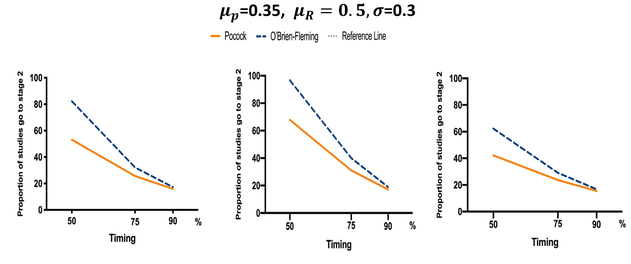
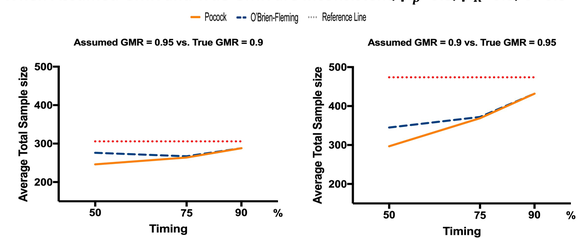


Figure 6 Average Total Sample Size by Timing of Interim Analysis

When Assumed GMR and True GMR are inconsistent, μ_p=0.3, μ_R=0.5, σ=0.3



Conclusion

Our simulation results show that

- All approaches control the Type I error rate under a negligible level.
- All approaches attain reasonable power.
- All approaches have minimal bias.
- Most importantly, the sequential comparative clinical endpoint BE study generally reduces the average total sample size of a fixed study design regardless which method is used.
- When the initial (fixed) study design sample size is large, 50% timing is generally more desired which gives a smaller sample size on average than 75% and 90% timing; when the initial sample size is small, 75% timing of interim analysis may be more desired which generally gives a smaller sample size than 50% and 90% timing. However, detailed scenarios may need to be evaluated on an individual basis.
- Between Pocock and O'Brien-Fleming alpha-spending approaches, Pocock saves more samples than O'Brien-Fleming in general due to the less stringent criteria for the interim analysis. On the other hand, O'Brien-Fleming alpha-spending approach has a higher power with a larger sample size.

IMPACT ON GDUFA

Our novel research in sequential comparative clinical endpoint BE study will

- Help generic drug applicants cut cost greatly and make clinical endpoint BE studies more affordable
- Hence improve the accessibility of generic drugs to the public.

DISCLAIMER

The opinions and information in this poster are those of the authors, and do not represent the views and/or policies of the U.S. Food and Drug Administration