Comparison of Different Methods of Sample Size Re-estimation for Therapeutic Equivalence (TE) Studies Protecting the Overall Type 1 Error

by
Diane Potvin

ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop / September 14th 2018

Outline

1. Therapeutic Equivalence Designs
2. Objectives
3. Simulations Basis
4. Method 1: Blinded Sample Size Re-estimation Method
5. Method 2: Sample Size Adaptive Sequential Design
6. Conclusions

1. Therapeutic Equivalence (TE) Designs

- Based on clinical endpoint
- Parallel Design: Test (T), Reference (R) and Placebo (P)
- Equivalence between Test and Reference (T/R and 90% CI between 80-125%) on PP.
- Superiority of Test and Reference vs. Placebo on mITT.
- Clinical endpoints used in TE often highly variable
- Require large sample size
- Major challenge: to obtain an accurate estimate of CV%

2. Objectives

- To develop methods for carrying out sample size reassessment at interim analyses without affecting the Type 1 error rate.
- Focus on continuous endpoints (e.g. change-from-baseline in FEV1).
- In these instances, the superiority criteria are generally easily met and do not drive sample size requirements.
- Focus on re-estimating sample size for equivalence at interim analyses

3. Simulations Basis - Parameters

- Normally distributed endpoint
- 1000 simulations per scenario
- CV% = 50, 60, 70, 80, 90 and 100
- Ratio ($\mu_{\Delta T} / \mu_{\Delta R}$) = 0.79, 0.90 and 0.95
- Superiority of T, R vs. P; non-superiority of T, R vs. P
- Sample size using a *4:4:1 ratio* (based on CV of 80%)
  - 1017 (452:452:113) initially planned patients in mITT
  - 918 (408:408:102) in PP
  - Sample size re-estimation at interim analyses (25% and/or 50%) to target 90% power on equivalence.
- Futility rule (stop if CP < 10% at interim)
- Max sample size of 1566 or 1840 (CV of 100%)

3. Simulations Basis – Initial Values

Table 1: Initial values for Simulations

Parameter	Setup value
μ_R	2.69L
μ_B	2.35L
$\mu_{\Delta R} = \mu_R - \mu_B$	0.34 L
$k = \mu_{\Delta T} / \mu_{\Delta R}$	0.95
$\sigma_{\Delta T,R}$	CV x $\mu_{\Delta R}$
$\sigma_{\Delta P}$	1.40 X $\sigma_{\Delta T,R}$
σ_{wi}	$\sigma_{\Delta i} / \sqrt{2}$
σ_b*	0.60L
δ	-0.017

3. Simulations Basis – Statistical Methods

- Assumptions:

Equivalence (per-protocol (PP)):

$H_{01}: \mu_{\Delta T}/ \mu_{\Delta R} < 0.80$ or $H_{02}: \mu_{\Delta T}/ \mu_{\Delta R} > 1.25$

$H_1: 0.80 \leq \mu_{\Delta T}/ \mu_{\Delta R} \leq 1.25$

Superiority (modified intent-to-treat (mITT)):

$H_0: \mu_i = \mu_P$ where i = R or T

$H_1: \mu_i \neq \mu_P$

- TE: ANCOVA with Treatment (T, R) as fixed effect and baseline as a covariate. (1-2α)% CI for the (T/R) ratio calculated using Fieller's theorem.
- Superiority: **T**est and **R**eference should both be statistically superior to **P**lacebo ($p<0.05$, 2-sided).
  - Similar ANCOVA model as described above, for both pairs of compared treatments.

4. Method 1 – Blinded Sample Size Re-estimation

- No possibility to conclude TE at interim
- Blinded estimate of sample size
  - Based on T, R combined
  - Based on Equivalence only

- Gould (1992), blinded estimate of variance:

$$\hat{\sigma}^2 \approx \frac{n-1}{n-2}\left(s^2 - {}^{\delta^2}/_{4}\right)$$

n: sample size (of both T and R combined)
S: sample SD estimate
δ: difference in means of T minus R under TE assumption

4. Method 1 – Blinded Sample Size Re-estimation

- Potvin (2017) Blinded estimates of mean R and CV :

$$\bar{X}_R \approx \bar{X}/(0.5 + 0.5\mathrm{k})$$

$$\mathrm{CV} = {}^{\hat{\sigma}}/_{\bar{X}_R}$$

k: T/R ratio under TE assumption

$\bar{X}$: overall mean of T, R combined

$\hat{\sigma}$: blinded estimate of variance (Gould)

4. Method 1 – Blinded Sample Size Re-estimation

- Interim Analyses
  - 25% interim: only increase allowed.
    - Trigger for increase (N estimated / N initially planned) of 1.05, 1.15 and 1.30.
  - 50% interim: increase/decrease allowed.
    - Trigger for increase/decrease of 1.05/0.95, 1.15/0.85 and 1.3/0.70.

4. Method 1 – Blinded Sample Size Re-estimation

Table 2: Overall type 1 error / Power (%) for therapeutic equivalence using blinded sample size re-estimation

CV%	Trigger 1.05 / 0.95	Trigger 1.15 / 0.85	Trigger 1.30 / 0.70
50	2.9 / 95.4	2.4/ 95.9	3.5 / 95.5
60	3.1 / 93.2	3.2 / 91.2	3.1 / 91.7
70	2.7 / 91.0	2.1 / 92.0	3.3 / 94.3
80	3.2 / 91.2	3.0 / 91.6	3.7 / 93.3
90	2.7 / 90.4	3.0 / 92.2	3.3 / 90.8
100	3.4 / 90.1	4.0 / 91.0	3.4 / 91.5

*Results presented for Superiority, $\mu_{\Delta T}/\mu_{\Delta R} = 0.79$ (non-equivalence scenarios to evaluate Type 1 error), $\mu_{\Delta T}/\mu_{\Delta R} = 0.95$ (equivalence scenarios to evaluate Power)

4. Method 1 – Blinded Sample Size Re-estimation

- Protects the overall Type 1 error for TE and superiority testing to a maximum of 5%
- Present an acceptable power.
- When the T to R ratio was less than 0.95 (i.e. a ratio of 0.90), the power to show TE was lower than 90%.

4. Method 1 – Blinded Sample Size Re-estimation

Table 3: Sample Size increases/decreases at 25%, 50% interim analyses

$\mu_{\Delta T}/\mu_{\Delta R}$	Trigger %	Interim	Action	%Studies	N (min, max)	Blinded CV	%Studies	N (min, max)	Blinded CV
0.79	105/95	25%	Increase	0.2	918 (918, 981)	0.67	80.4	1116 (918, 1413)	0.88
		50%	Decrease	99.9	633 (477, 918)	0.67	38.0	1092 (801, 1413)	0.88
			Increase	0.0			25.8		
	115/85	25%	Increase	0.0	918 (918, 918)	0.67	59.8	1098 (918, 1413)	0.89
		50%	Decrease	98.3	633 (459, 918)	0.67	9.2	1103 (891, 1413)	0.88
			Increase	0.0			13.6		
	130/70	25%	Increase	0.0	918 (918, 918)	0.67	33.3	1047 (918, 1413)	0.89
		50%	Decrease	64.3	710 (477, 918)	0.66	0.1	1071 (918, 1413)	0.89
			Increase	0.0			7.6		
0.9	105/95	25%	Increase	0.0	918 (918, 918)	0.62	50.7	1007 (918, 1413)	0.83
		50%	Decrease	100.0	550 (459, 738)	0.62	47.5	964 (693, 1305)	0.82
			Increase	0.0			15.4		
	115/85	25%	Increase	0.0	918 (918, 918)	0.62	25.4	983 (918, 1413)	0.83
		50%	Decrease	100.0	549 (459, 693)	0.62	8.2	975 (675, 1386)	0.82
			Increase	0.0			6.4		
	130/70	25%	Increase	0.0	918 (918, 918)	0.62	9.3	952 (918, 1413)	0.83
		50%	Decrease	96.2	558 (459, 918)	0.62	0.0	954 (918, 1413)	0.82
			Increase	0.0			0.7		
0.95	105/95	25%	Increase	0.0	918 (918, 918)	0.60	34.9	975 (918, 1413)	0.80
		50%	Decrease	100.0	520 (459, 693)	0.60	57.5	915 (675, 1224)	0.80
			Increase	0.0			11.0		
	115/85	25%	Increase	0.0	918 (918, 918)	0.60	14.7	952 (918, 1413)	0.80
		50%	Decrease	100.0	519 (459, 657)	0.60	11.2	933 (639, 1278)	0.80
			Increase	0.0			2.6		
	130/70	25%	Increase	0.0	918 (918, 918)	0.60	3.7	931 (918, 1413)	0.80
		50%	Decrease	99.4	519 (459, 918)	0.60	0.1	932 (639, 1413)	0.80
			Increase	0.0			0.3		

*CV=60% (highlighted in green), CV=80% (highlighted in blue)

4. Method 1 – Blinded Sample Size Re-estimation

- As CVs increase, % studies requiring a sample size increase at 25% and at 50% also increases.
- Larger triggers 130/70% preferable in avoiding too many sample size adjustment at interim.
- For ratios of 0.79 and 0.90, the method tends to increase sample size.
  - Blinded estimate of CVs provides larger values than expected, since the initial assumption of a ratio of 0.95 is not true.

5. Method 2: Sample Size Adaptive Sequential Design

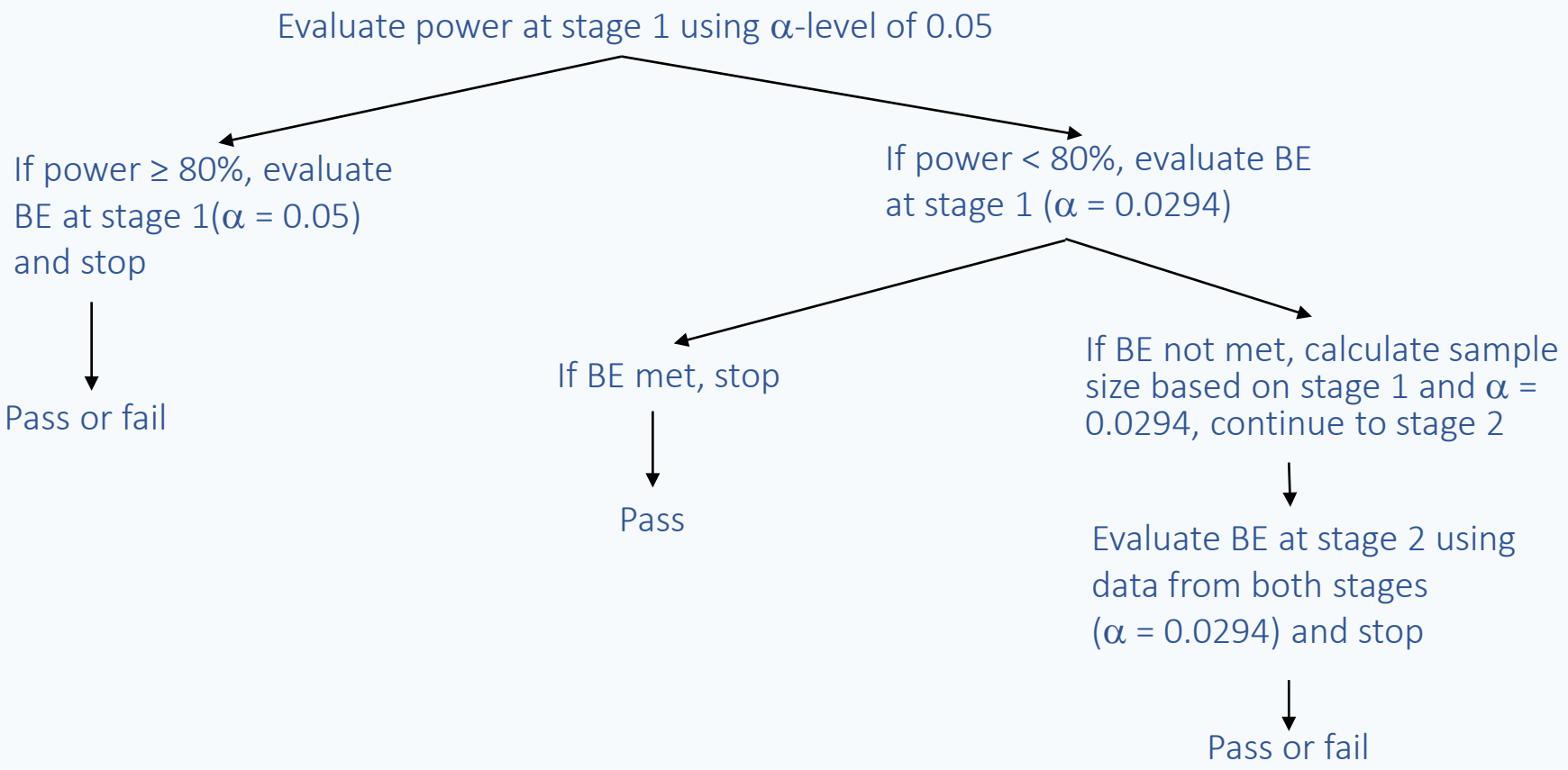
- Interim analysis at 50%
- Determine the variance estimate
- Re-adjust sample size if required
- Possibility to stop at interim for equivalence
- Group sequential designs and sample size re-estimation methods:
  - Not validated for parallel studies with Normally distributed data and two one-sided t-tests on T/R ratio using Fieller's theorem.

5. Method 2: Sample Size Adaptive Sequential Design

- The adaptation only done on the equivalence part
  - Decision to stop at interim
  - Type 1 error adjustment
  - Sample size increase
- Superiority testing follows decision's process triggered by equivalence
  - Only tested once (α=5%)
- Overall Type 1 error and power calculated for both equivalence and superiority combined

5. Method 2: Sample Size Adaptive Sequential Design

- Adaptive sample size sequential design based on Pocock (Method C)



5. Method 2: Sample Size Adaptive Sequential Design

Table 4. TE results using Potvin's method C for Equivalence

%CV	Overall Type 1 error (%)	Overall Power (%)	N Avg T + R (PP) / H_0	N Avg T + R (PP) / H_1	N T+R single stage design
50	2.7	95.7	597.64	411.26	320
60	2.3	95.9	857.66	455.25	456
70	3.1	95.7	1161.74	565.44	616
80	3.0	97.7	1378.45	768.36	808
90	3.4	97.5	1444.76	1001.88	1024
100	3.0	93.1	1455.84	1208.55	1256

5. Method 2: Sample Size Adaptive Sequential Design

- Protect overall Type 1 error
- Power > 90% : due to simulations initial settings (use the larger CV between T and R)
- Optimal sample size under H1
- Large sample size under H0 : would need a different futility rule.

6. Conclusions

- Method 1
  - Very interesting and novel solution for TE trials for which there is uncertainty in the initial sample size estimate.
  - Interim analyses can be performed throughout the trial in order to adjust the sample size (increase or decrease) in a blinded fashion
  - No need to adjusting the Type 1 error for interim looks, since blinded estimates of CVs are being performed.
  - Prevent the inflation of the Type 1 error
- Method 2
  - Allows to stop at interim and conclude TE
  - Protect overall Type 1 error with adequate power
  - Would need additional R&D
    - Dichotomic endpoints
    - Endpoints with strong placebo effect (superiority is the limiting factor)
    - More efficient futility rule (to stop early under H0)

References

- Gould A.L., Weichung J.S. Sample size re-estimation without unblinding for normally distributed outcomes with unknown variance. *Commun. Statist. – Theory Meth*., 21(10), 2833-2853 (1992).
- Hauschke D., Kieser M., Diletti E. and Burke M. Sample size determination for proving equivalence based on the ratio of two means for normally distributed data. *Statist. Med*. 18, 93-105 (1999).
- Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ, and Smith RA. Sequential Design Approaches for Bioequivalence Studies with Crossover Designs. Pharmaceutical Statistics 2008; 7:245-262.
- Potvin D., Briand O., Morin J., Lau H., Valente A., Patterson B. (2017, November). Novel Approach for Blinded Sample Size Re-Estimation for Therapeutic Equivalence (TE) Studies Protecting the Overall Type 1 Error. Poster session presented at the American Association of Pharmaceutical Scientists Annual Meeting, San Diego, CA.

Thank you

- Olivier Briand (Senior Data Manager and SAS Programmer)
- Josée Morin (Sr. Biostatistician)
- Céline Le Naour (SAS Programmer)