# A bioequivalence study design in Immunology which includes the option for sample size re-estimation (SSR) at the Interim Analysis

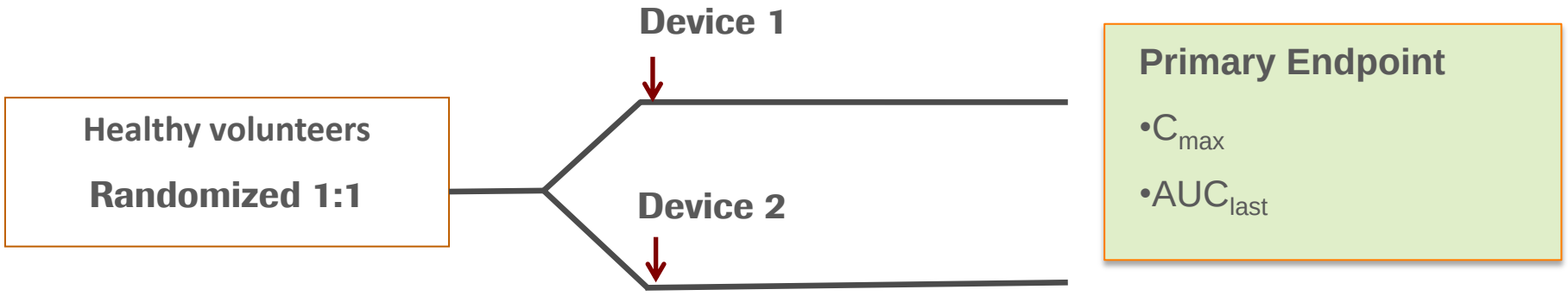
*Jennifer Pulley*
*(Statistical Scientist)*

# Overview

- Study design
- Interactions with the Health Authorities
- Revised study design
- Sample size re-estimation (SSR) algorithm and analysis approach

# Bio-equivalence Study Design

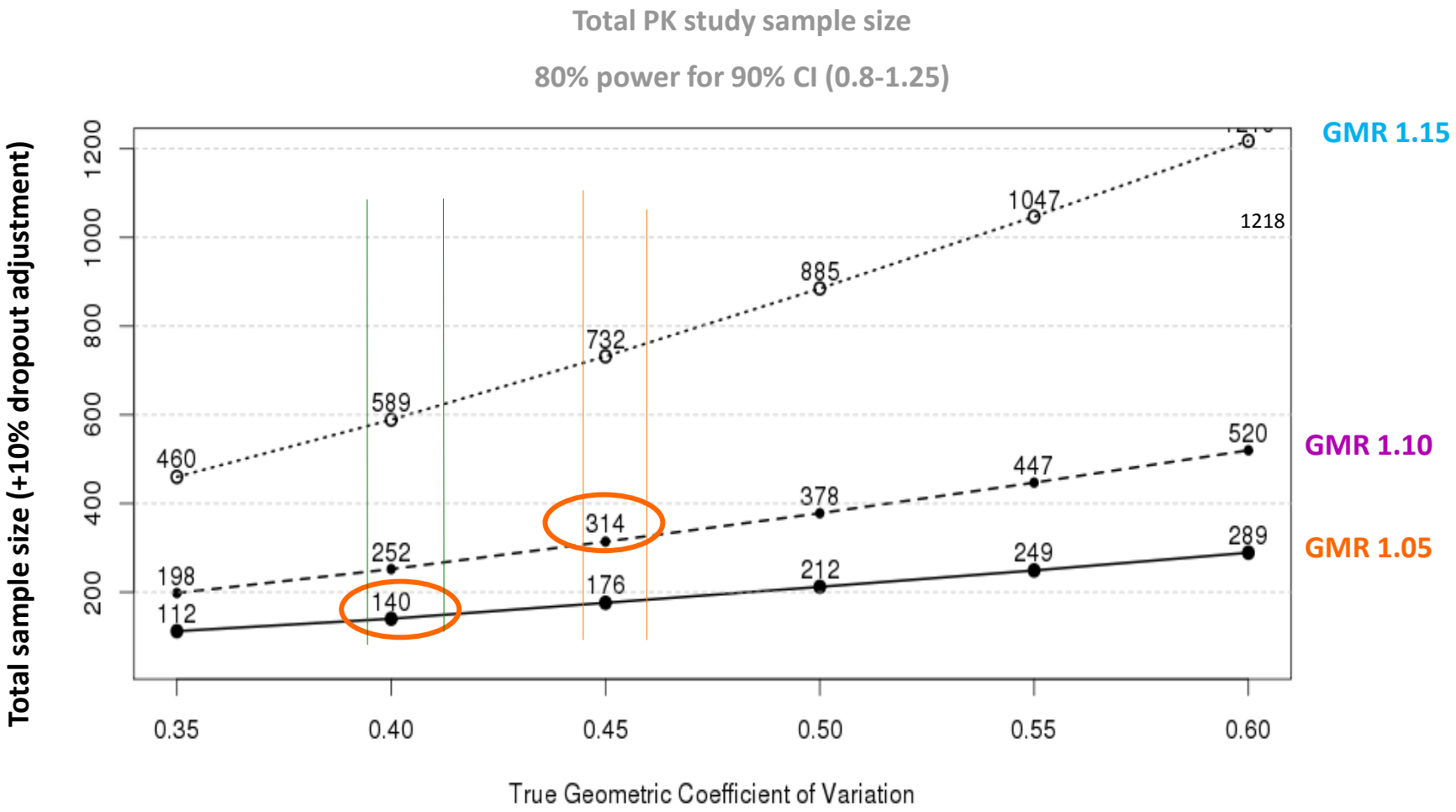
- Support device development for a large phase III immunology molecule
- Drug has a long half life which lead to use of a parallel design over a cross over design




**Bioequivalence/comparability criterion:**

- 90% confidence interval for geometric mean ratio (GMR) (Device 1 vs. Device 2) of both Cmax and AUClast is contained within 0.8–1.25

# Sample Size Options for varying GMR and CV

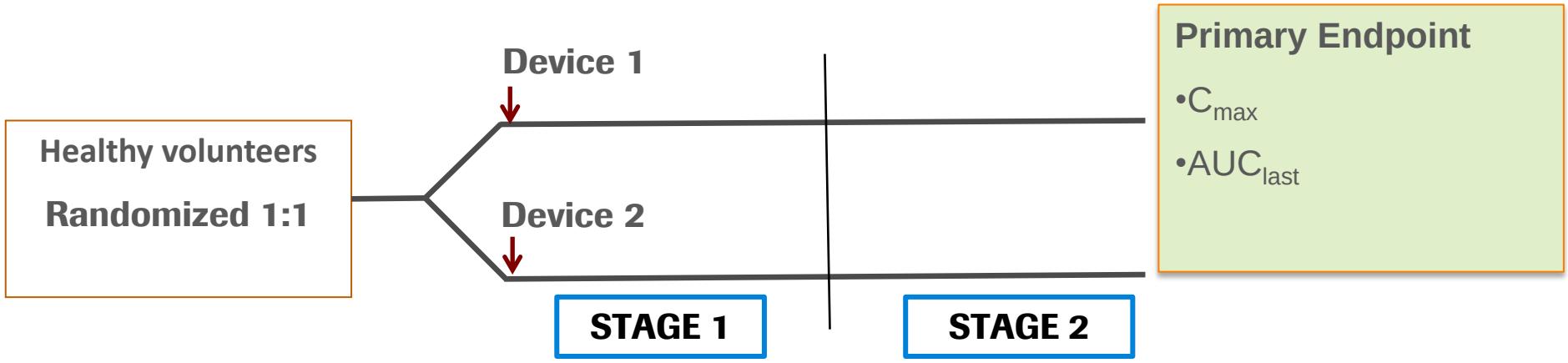



GMR=Geometric mean ratio (ratio of mean AUC or Cmax for the two devices)

Plot created using R

Power.Tost package in R was used to calculate the Sample size calculation

# PK Comparability Interim Analysis Study Design



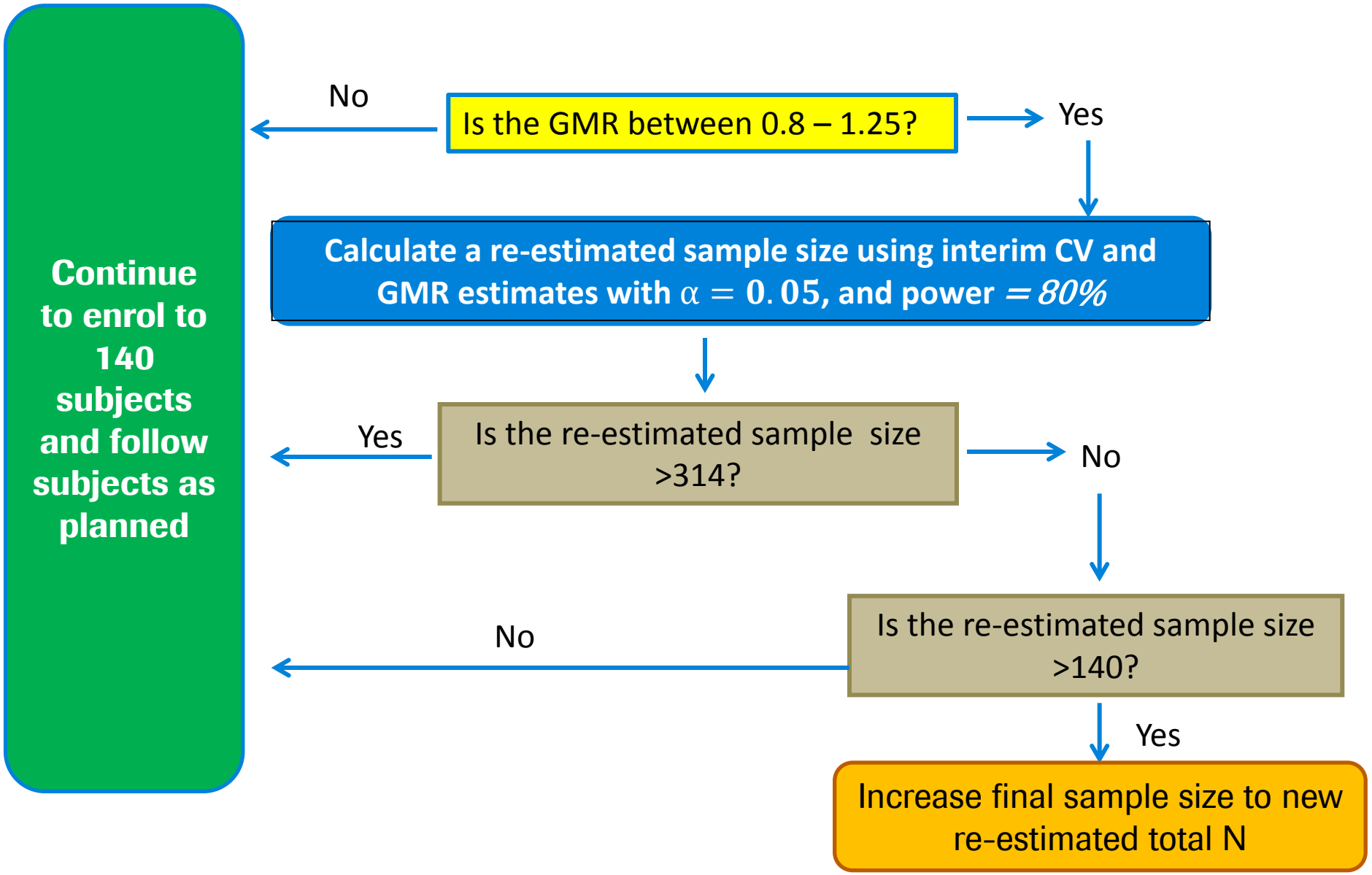

**Interim Analysis – Sample Size Re-estimation (SSR)**

Following stage 1 and once 60 subjects (30 per arm) provided PK samples out to day X the IA will be conducted by a third party vendor to re-estimate sample size

**Bioequivalence/comparability criterion**

90% confidence interval for geometric mean ratio (GMR) (Device 1 vs. Device 2) of both Cmax and AUClast is contained within 0.8–1.25

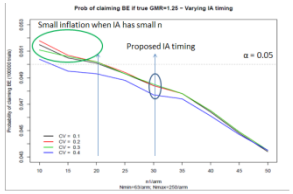
# Two Stage BE Study Design – Decision Criteria



# Rational for Interim Analysis

- Allow team to get a accurate estimate of GMR and CV assumptions and adjust the sample size if required
- IA study design had been conducted in other disease areas within Roche
- Timelines
- Adjustment was not made for Type I error due to BE not being assessed at the IA
  - EMEA 2012 paper discussing the revised EMA guideline (EMA author)*:* "*The plan to spend alpha must be pre-defined in the protocol. […]. It is also possible to distribute the alpha differently, and as an extreme case,* ***it is acceptable to plan no alpha expenditure in the interim analysis when it is designed to obtain information on formulation differences and intra-subject variability and 90% CI are not estimated at the interim stage***."
- FDA – no guidance on SSR for BE studies.
- Plan to keep the design straightforward

# Interactions with FDA




Roche

FDA

**Conclusion**

Post TC with FDA, team needed to either revise analysis plan adjusting test statistic/alpha levels, or provide a theoretical argument for not performing a type I error or test statistic adjustment (i.e. simulations alone not accepted by the FDA).

# Revised SSR Study Design Options*

| Design | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Power Method | Conditional Power (CP) | | Unconditional Power (UCP) | |
| Adjustment | Adjusted | Unadjusted | Adjusted | Unadjusted |
| Confidence Interval Method | Exact 2-stage based | Standard | Exact 2-stage based | Standard |

- CP method allows the probability to reject Ho given the stage 1 data and this is based on Promising Zone approach.
- UCP method, allowed the sample size to be re-estimated based on assumptions from stage 1 and assuming set power e.g 80%.
- The two methods allowing for adjustment use adjusted CI to control type I error
- An equivalent 90% confidence interval with 'adjusted' confidence bounds was derived, which can be used in the usual way to assess the BE criteria for AUC/Cmax.

*in collaboration with Cytel, Cambridge USA

# The 'Promising Zone' Approach for SSR

Method first described by Mehta and Pocock (2011):

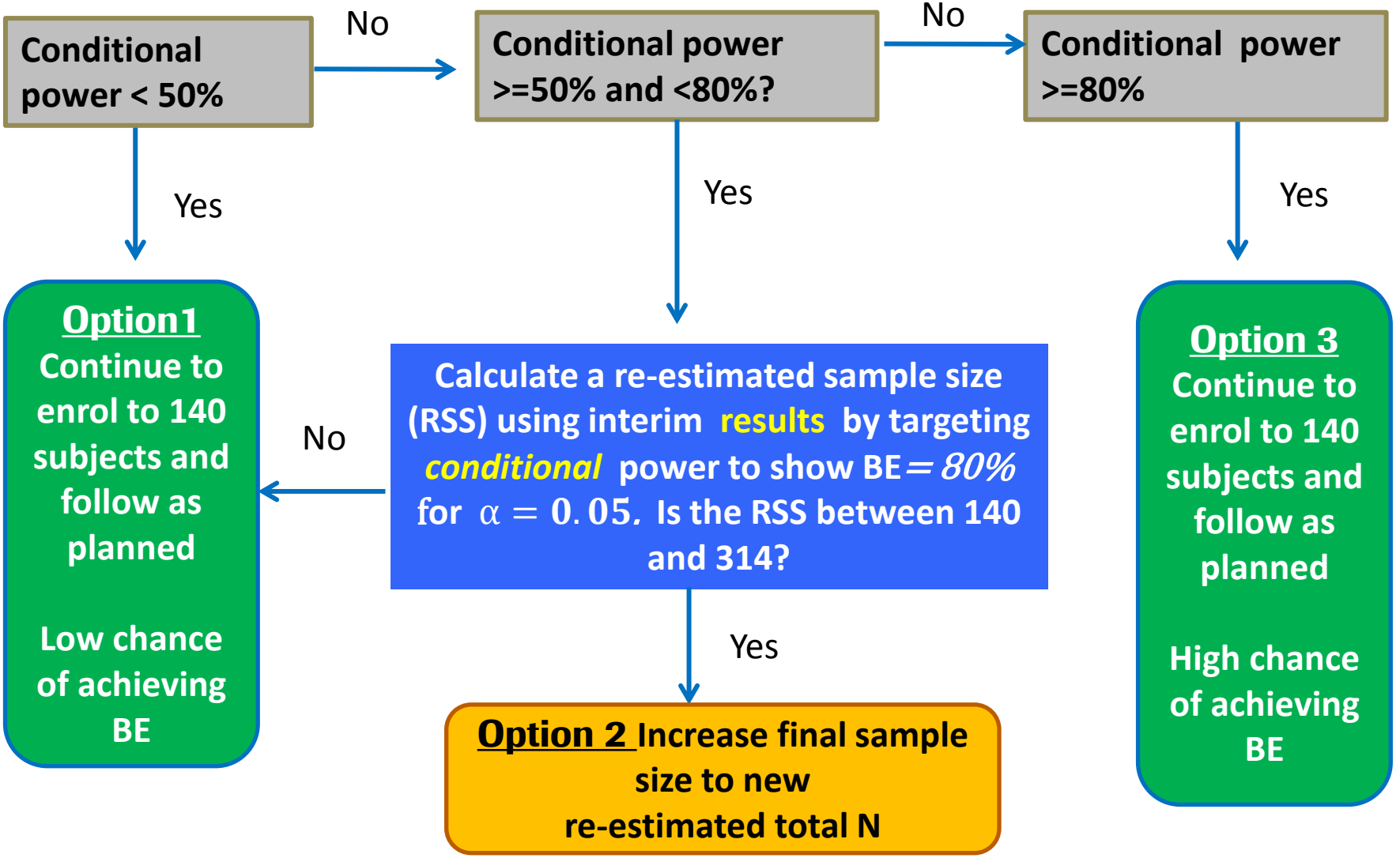
**Option1**: If the conditional power (CP) falls below a predefined boundary $CP_{min}$, then the interim results are considered unfavorable and study continues with the originally planned second stage sample size

**Option 2:** If CP ≥ $CP_{min}$ but ≤ the planned power $1-\beta$, then the results are seen as promising. For these interim results, the sample size for the second stage is increased such that CP ≥ $1-\beta$.

**Option 3:** If the conditional power is ≥ $1-\beta$, the results are seen as favorable and the originally planned second stage sample size is maintained.

It's been previously shown (Chen et al) that type I error is not increased if $CP_{min} \geq 50\%$

# Updated Proposed Two Stage BE Study Design




Note: the 50-80% range is also called a 'promising zone' for CP

CP Min = 50%, $1-\beta=80\%$

# Testing for Bio-equivalence

- Following the SSR step stage 2 of the study is completed and stage 1 and stage 2 data is combined to test for bio-equivalence between the two devices for the two PK parameters Cmax and AUClast
- PK parameters log(Cmax) and log (AUClast) assumed to be independent and normally distributed
- For Bioequivalence test two null hypotheses and one alternative hypothesis where $\delta_L = \log(0.8)$ & $\delta_U = \log(1.25)$ and alpha=5%
  - $H_0: \delta \leq \delta_L$ or $\delta \geq \delta_U$ against $H_a: \delta_L < \delta < \delta_U$
- Alternatively this testing can be done as two one-sided non-inferiority tests
  - $H_{01}: \delta \leq \delta_L$ against $H_{a1}: \delta > \delta_L$
  - $H_{02}: \delta \geq \delta_U$ against $H_{a2}: \delta < \delta_U$

For each one-sided test a separate test statistic is defined

# CHW Test Statistic

- Cui-Hung-Wang (CHW) approach defines a weighted statistic ($Z^*$), data from the 2-stages by weighing the stage wise Wald test statistics:

- Wald Test Statistic is defined as

Stage 1: $Z_{1,L} = (\hat{\delta}_1 - \delta_L)/SE(\hat{\delta}_1)$ , $Z_{1,U} = (\hat{\delta}_1 - \delta_U)/SE(\hat{\delta}_1)$

Stage 2: $Z_{2,L} = (\hat{\delta}_2 - \delta_L)/SE(\hat{\delta}_2)$ $Z_{2,U} = (\hat{\delta}_2 - \delta_U)/SE(\hat{\delta}_2)$ ,

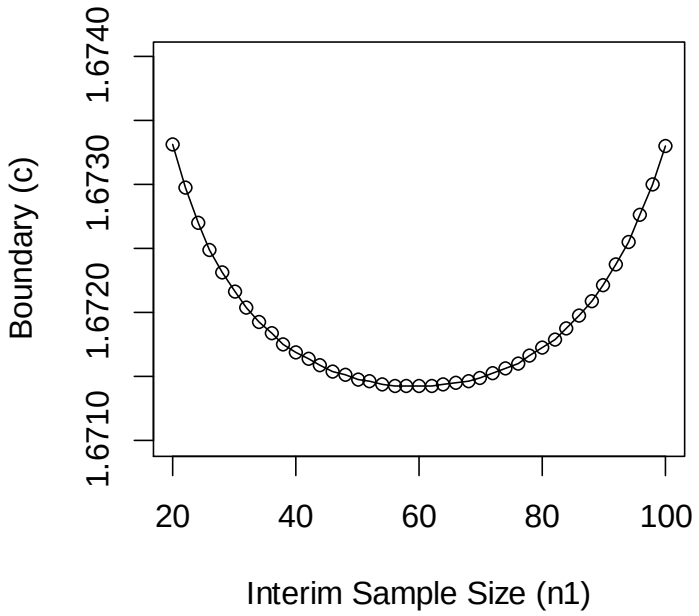
where $\hat{\delta}_1$ = difference in group means in stage 1, $\hat{\delta}_2$ = difference in group means in stage 2, and $SE(\hat{\delta}_1)$ and $SE(\hat{\delta}_2)$ are the corresponding standard errors for each stage

- The CHW statistic combines the two stages statistics together and applies the weighting

$$Z_L^* = \sqrt{W_1}Z_{1,L} + \sqrt{W_2}Z_{2,L} \quad \text{and} \quad Z_U^* = \sqrt{W_1}Z_{1,U} + \sqrt{W_2}Z_{2,U}$$

# Adjusted Critical Values and Adjusted CI

- Calculated adjusted critical values (c) based on the t-distribution (to avoid type I error inflation due to small sample sizes) depend only on $n_1$, $n_2$, $N_{max}$ and α.
- An adjusted two-sided CI, ($C_L$, $C_U$), replacing the conventional 90% CI, and can be defined equivalent to the pair of adjusted tests, which was shown to have expected coverage, i.e. Prob($C_L$ < δ < $C_U$) ≥ 1- 2α.
- BE would be claimed if both $Z_L^*$ > c and $Z_U^*$ < -c or equivalently, if ($C_L$, $C_U$) is completely contained in the (0.8, 1.25) bioequivalence range.
- Figure illustrates calculated critical values for $n_1$ = 20 - 100, $n_2$ = 126 – $n_1$, $N_{max}$ = 286, α = 0.05




n1= sample size stage 1
n2 = planned sample size stage 2
Nmax= actual sample size of stage 2

# Summary

- Team revised the study design to use a method for analysing the stage 1 and stage 2 data which would not inflate type I error
- Type I error is built into an adjustment in the 90% Confidence Intervals around the parameter estimates (via adjusted critical values for combined stage 1 and 2 test statistics)
- Team submitted the revised study design 1 to the FDA which provided a proof for control of the type 1 error using the t-distribution

| Design | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Power Method | Conditional Power (CP) | | Unconditional Power (UCP) | |
| Adjustment | Adjusted | Unadjusted | Adjusted | Unadjusted |
| Confidence Interval Method | Exact 2-stage based | Standard | Exact 2-stage based | Standard |

# Follow-up with FDA

- Did the FDA support the use of the revised study design allowing control for type I error?

# Follow-up with FDA

- Did the FDA support the use of the revised study design allowing control for type I error?

- *FDA acknowledge the statistical approach adequately addressed the control of type 1 error.*
- *However, as a matter of regulatory policy, we have not accepted the use of such an approach in support of bioequivalence/comparability studies. Implementing your proposed approach will constitute establishing a new policy for which we are required to follow proper procedures.*

# References


- Hsiao.S, Lingyun.L (2016). Unblinded Sample Size Re-Estimation in Bioequivalence Trials with Small Sample Sizes*:* Joint Statistical Meeting July 30-Aug 4, Chicago

- Mehta CR, Pocock SJ (2011). Adaptive increase in sample size when interim results are promising: A practical guide with examples. *Statistics in Medicine,* **30**:3267–3284

- Chen, Y. H. J., DeMets, D. L. and Gordon Lan, K. K. (2004), Increasing the sample size when the unblinded interim result is promising. Statist. Med., 23: 1023–1038.

*Doing now what patients need next*