

## SAMPLE SIZE RECOMMENDATION FOR A BIOEQUIVALENT STUDY

Chandrasekhar Bhupathi <sup>1</sup>

*Department of Statistics, O U Science College, Osmania University, Hyderabad, India.*

Venkata HaraGopal Vajjha

*Department of Statistics, O U Science College, Osmania University, Hyderabad, India.*

### 1. INTRODUCTION

In clinical trials, as part of drug development process, we may need to compare the experimental drug or treatment or formulation with the existing treatment to check the drug BA/BE properties.

In this section, we look into the general definitions of BE, its importance in clinical trials and the criterion to accept or reject the BE. Also, we cover the BE guidelines and brief literature review.

#### 1.1. Bioequivalence

If two products are said to be *bioequivalent* it means that they would be expected to be, for all intents and purposes, the same. As per “International Council for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use” (ICH) guideline document BE is defined as,

*‘the absence of a significant difference in the rate and extent to which the active ingredient or active moiety in pharmaceutical equivalents or pharmaceutical alternatives becomes available at the site of the drug action when administered at the same molar dose under similar conditions in an appropriately designed study’*

##### 1.1.1. Importance in clinical Trials

The objective of a BE is comparison of two products. The utilization of generic drug is growing rapidly. So here a BE study is a must to compare the branded drug with generic drug and need to prove that the biological properties of these two drugs are equal. In BE, the comparison is not necessarily between generic and branded. It can be between two formulations of same drug, like, capsule vs tablet and etc.

---

<sup>1</sup> Corresponding Author. E-mail: chandu.bhupathi@gmail.com

### 1.1.2. BE acceptance criterion

To assess the BE, we need the pharmacokinetic (PK) parameters - Cmax and AUC. In general, these are referred as the primary PK parameters in the clinical trial. These are derived using the plasma concentrations, which are derived using the blood collected at specified intervals. The general criterion to accept that the drugs are bioequivalent is,

‘90% CI of the geometric mean ratios of Cmax and AUC between test and reference fall within the acceptance range 80% to 125%’ [Shen M et al. (2014); BPJ (2009); Tsai CA et al. (2014)]

## 2. LITERATURE REVIEW

There are clear guidances on the sample size and power calculation. Even health authorities provide suggestions on required sample size and power for Bioequivalence (BE) studies. But, when we compare the clinical trials conducted on healthy volunteers (HV) or subjects, especially for BE studies, with the guidelines, we are happy that clinical trials are being conducted according to the guidelines provided by health authorities. However, guidelines talk about the minimum required sample size with desirable power.

### 2.1. Recommended sample size limits for a BE study

As per the guidelines provided by Europe, Canada, New Zealand and USA (a pilot study) the sample size should be a minimum of 12 subjects. And, for Japan it should be a sufficient number;

But for the maximum limit, this is not clearly mentioned or specified. This will be judged by Institutional Review Board (IRB)/Institutional Ethics Committee (IEC) or local authorities. As per section 3.5 of ICH E9, the number of subjects in a clinical trial should always be large enough to provide a reliable answer to questions addressed.

### 2.2. What is lacking?

The guidelines say that, for a BE study, the sample size should be large enough to provide reliable results. However guidelines talk about the minimum required sample size with desirable power. But, **how large** is sufficient is not clear from these guidelines provided by health authorities.

From one of the article [Davit BM et al. (2009)], based on 12 year data on generic drugs in US FDA [FDA guidance (2011), (2014)], the number of participants range from 12 to 172, but most studies have enrolled subjects from 24 to 36. We didn't have information about how many studies have been planned with reasonable number of subjects.

So the question is, **why don't we stick to this least power, suggested by HA, that a BE study requires?** That means, instead of aiming for greater than 95% power and recruiting more HV into the study, can we reduce the sample size to satisfy the sufficient power, but still prove or disprove the BE? If we do so, what are the factors that will get affected with such change?

In this paper, we compare the statistical analysis results with different sample sizes and power. We also touched upon several factors affecting with the sample size in drug development.

In the examples, we used the simulated data and results are based on simulated data;

### 3. EVALUATION

It has been observed that most of the Bio equivalence (BE) studies are being planned and conducted with more sample size than required. As we all are aware that time and money are directly proportional to sample size, it has direct impact on pricing of the drug, which has been observed more carefully in recent times by health authorities (HA) and pharmaceutical companies.

We all know that sample size calculation for a BE study is dependent on multiple factors like power, intra subject coefficient of variation, expected geometric mean ratio and etc. Also, as we know that power and sample size are proportionately related, if we increase the power the sample size increases and vice-versa. Generally we use a fixed number as a power in the sample size calculation. So, if we now focus on what could be a reasonable statistical power to conduct a BE study, perhaps we may get more closer to the details of the problem.

As per the health authorities, we need to plan the study with at least 80% power. This doesn't mean that we should plan with only 80% or 81% power. It also doesn't mean that we can plan the study with a power of 95% or above.

From one of the World Health Organization's training workshops presentations [Alfredo G, WHO (2009)], it is clear that there is no need to include too many (extra) subjects.

#### 3.1. Results

To assess this observation, we have used two different studies. The first study is planned with 54 subjects at 98% power. The result of this is that the drugs are bioequivalent. And the second one it is planned with 38 subjects at 94% power. The result of this study is that the drugs are not bioequivalent.

##### 3.1.1. Illustration One

Imagine if a particular BE study, a two period, two treatments crossover study is conducted with 54 subjects (98% power) to confirm bio equivalence of treatments. But, if we can conclude the same with 26 subjects (80% power, as per FDA guidelines), then with almost 50% of lesser sample size, we can conclude the same. We can save blood of excess subjects or these subjects can get enrolled into other studies.

Consider the below values as fixed for the sample size calculation.

Lower Bound (LL) = 0.80, Upper Bound (UL) = 1.25, Alpha = 0.05,  
Geo Mean Ratio (GMR) = 0.947, Coefficient of Variation (CV) = 0.239

The sample size and power vary like below,

Sample	54	50	47	44	35	30	26	24
Power	97.9	97.0	96.0	95.0	90.0	85.0	80.0	76.6

As per the above results, we can conclude that the drugs are bioequivalent with 26 subjects; almost 50% of lesser sample size.

The statistical model was repeated for different sample sizes (44, 40, 34, 30, 26), each with 100 simulations. The GMR, LL and UL are displayed in the below graph for better understanding.

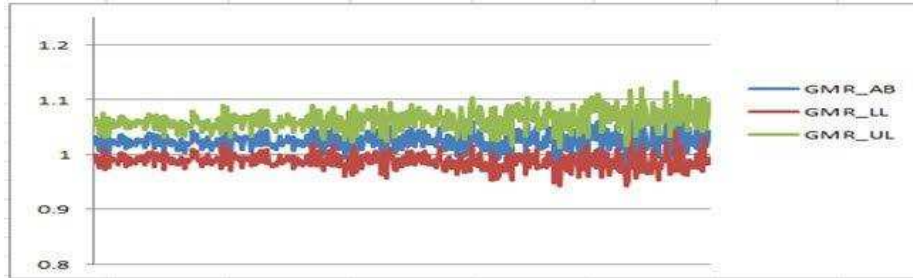


Figure 1 – Graphical representation of GMR, LL and UL

The GMR, LL and UL are within the BE acceptable limits, that is within 0.8 to 1.25 range.

Also, we tried to observe the sample size fluctuations with different GMR and Power combinations by keeping the other factors unchanged

Lower Bound = 0.80, Upper Bound = 1.25, Alpha = 0.05, Coefficient of Variation = 0.239

### 3.1.2. Illustration Two

For illustration, imagine if a particular BE study, a two period, two treatments crossover study is conducted with 38 subjects (93.8% power) to confirm no bioequivalence of treatments. But, if we can conclude the same with 26 subjects (82.7% power, as per FDA guidelines), then with almost 50% of lesser sample size, we can conclude the same. We can save blood of excess subjects or these subjects can get enrolled into other study.

Consider the below values as fixed for the sample size calculation.

TABLE 1  
Sample size variations for different Power and GMR values

Power	GMR											
	0.85	0.90	0.92	0.95	0.97	1.00	1.02	1.05	1.07	1.10	1.15	1.20
98	415	111	80	53	43	37	40	52	65	95	220	914
97	378	101	73	49	39	34	37	47	59	86	200	831
96	350	94	67	45	37	32	34	44	55	80	186	771
95	329	89	63	43	35	31	33	42	52	75	175	723
90	261	70	51	34	29	26	27	33	41	60	139	573
85	219	59	43	29	25	23	24	29	35	51	117	481

Lower Bound (LL) = 0.80, Upper Bound (UL) = 1.25, Alpha = 0.05,  
 Geo Mean Ratio (GMR) = 1.05, Coefficient of Variation (CV) = 0.237

The sample size and power varies like below,

Sample	42	40	38	34	30	28	26	24
Power	95.7	94.8	93.8	91.3	87.7	85.4	82.7	79.4

As per the above results, we can conclude that the drugs are not bioequivalent with 26 subjects; almost 30% of lesser sample size. The statistical model was repeated for different sample sizes (34, 30, 28, 26), each with 100 simulations. The GMR, LL and UL are displayed in the below graph for better understanding.

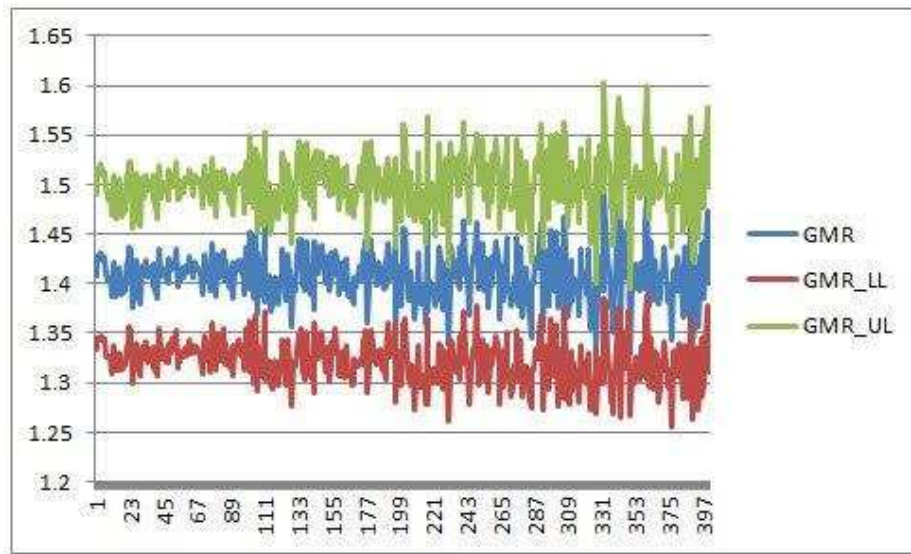


Figure 2 – Graphical representation of GMR, LL and UL

The GMR, LL and UL are above the BE acceptable limits, that is >1.25.

Below we tried to observe the sample size fluctuations with different GMR and Power combinations by keeping the other factors unchanged.

LL = 0.80, UL = 1.25, Alpha = 0.05, Coefficient of Variation = 0.237

### 3.2. Impact

It is clear from above results, that high power and sample size are not needed to prove or disprove a BE study. We explained the impact of high sample size on a trial and also on overall drug development below in brief.

As we all know that power and sample size are proportional to each other, more power leads to plan the study with more sample size. Once we plan the study with higher sample size we end up spending more time and money.

TABLE 2  
Sample size variations for different Power and GMR values

Power	GMR											
	0.85	0.90	0.92	0.95	0.97	1.00	1.02	1.05	1.07	1.10	1.15	1.20
95	324	87	62	42	34	30	32	41	51	74	172	712
94	306	83	59	40	33	29	31	39	48	70	163	673
91	267	72	52	35	29	26	28	34	42	61	142	586
88	238	65	46	32	27	24	25	31	38	55	127	523
85	216	59	42	29	25	23	24	28	34	50	115	473
83	203	55	40	27	24	22	23	27	32	47	108	445

Apart from time and money, there are other important factors such as blood, medical waste and participants that can be saved OR used more appropriately if we can plan BE studies with reasonable sample size.

Pricing of drug-As per one of the latest article the average branded prescription drug prices have nearly doubled over the last five years. As per Tufts CSDD latest report [Kim Jannsen (2016)], cost of developing new drug would be around \$1 to \$2.6 Billion.

Blood-unavailability of blood is the major cause of death during accidents. As per world blood bank some where someone needs blood for every 2 seconds. By recruiting more healthy volunteers, we are blocking them for a specific time. At the same time we are wasting their precious blood, that can be used to save others' lives.

Enrollment/participation into clinical trials (CT) is another issue the industry is facing. Several awareness programs are being conducted to educate and encourage people to participate in CT.

Biomedical waste can be reduced by recruiting only required number of subjects, which has huge impact on the environment.

#### 4. CONCLUDING REMARKS

This observation is based on simulated data for which the drugs are bioequivalent for one study and not bioequivalent for other study. Even with a small sample size we are able to prove it. Planning the study exactly at 80% power may not be suggestable. But, for sure, there is no need to aim for high power. Based on the simulated data, instead of high power, aiming 85% power also reduces a considerable amount of sample size; this can have an impact on 10 to 24 healthy volunteers. From the results obtained, we can state that a power of 85% would be reasonable for a BE study to conduct on healthy volunteers. This helps a lot in saving time, blood and money. Also, it is not ethical to involve more healthy volunteers than required and expose them to the drugs which are still under testing.

#### ACKNOWLEDGEMENTS

Special thanks to colleagues and friends for their valuable comments and suggestions. The content is solely the responsibility of the authors and does not necessarily represent the official views or opinions of Novartis.

## REFERENCES

- ALFREDO GARCIA ARIETA, PHD (2009). *Statistical considerations, Training workshop: Training of BE assessors*. World Health Organization.
- DAVIT BM, NWAKAMA PE, BUEHLER GH, CONNER DP, HAIDAR SH, PATEL DT, YANG Y, YU LX, WOODCOCK J (2009). *Comparing generic and innovator drugs: A review of 12 years of bioequivalence data from the united states food and drug administration*. Foundations of Statistical Inference, Ann Pharmacother, 43, no. 10, p. 158397.
- FDA GUIDANCE FOR INDUSTRY (2011). *Statistical Approaches to Establishing Bioequivalence*. Food & Drug Administration.
- FDA GUIDANCE FOR INDUSTRY (2014). *Bioavailability and Bioequivalence Studies Submitted in NDAs or INDs General Considerations*. Food & Drug Administration.
- KIM JANNSEN (2016). *Cost of namebrand prescription drugs doubled over last 5 years*. Report on ChicagoTribune.
- S. RANI, A. PARGAL (2004). *Bioequivalence: An overview of statistical concepts*. Indian J Pharmacol, 36, no. 4, p. 209216.
- SHEN M, RUSSEKCOHEN E, SLUD EV (2014). *Exact calculation of power and sample size in bioequivalence studies using two onesided tests*. Pharmaceutical Statistics, Wiley Online Library, 14, no. 2, p. 95101.
- SPECIAL EDITION GENERICS (2009). *What is Bioavailability and Bioequivalence?* Best Practice Journal (BPJ).
- TSAI CA, HUANG CY, LIU JP (2014). *An approximate approach to sample size determination in bioequivalence testing with multiple pharmacokinetic responses*. Statistics in Medicine, Wiley Online Library, 33, no. 19, p. 33003317.
- WORLD BLOOD BANK (2016). *World Blood Bank website*. URL <http://www.worldbloodbank.org/>.

## SUMMARY

There are clear guidelines and suggestions on the sample size and power calculation from health authorities (HA) for Bio equivalence (BE) studies in Healthy volunteers (HV). The suggested power is at least 80% and type 1 error is 5%. In real life situations, the clinical trials plan with more than 80%, giving rise to larger sample size. The increased power means more subjects, more wastage of time and more resources to complete the study, resulting in more money spent. This paper attempts to show how much reduction in the sample size can be achieved without affecting the scientific validity of the study and also the brief summary on the overall effect of reduced sample size on resources (subjects, time, blood and cost). We executed simulations in order to show the impact on the power and the 2 one sided confidence interval approach to show the study equivalence or otherwise. For illustration purpose, a couple of 2 period cross over studies were considered. 100 simulations were executed with different sample sizes to compare with the original results.

**Keywords:** Bioequivalence; Healthy volunteers; Sample Size; Power.