Use of Geometric Mean in Bioequivalence Trials

Abdur Rasheed¹,* and Afaq Ahmed Siddiqui²

¹Department of Research, Dow University of Health Sciences, Karachi, Pakistan
²Faculty of Pharmacy, University of Karachi, Pakistan

Abstract: Bioequivalence data often do not follow the normality assumption on the linear (original) scale, therefore in that situation, the use of the logarithmic transformation is recommended. In the bioequivalence analysis, confusion arises about the use of geometric mean ratio when the logarithmic transformation is recommended by the regulatory authorities. The purpose of this research paper is to clear this confusion. Different average bioequivalence criteria are also reviewed in this paper.

Keywords: Geometric mean, Average bioequivalence, logarithmic transformation, Bioequivalence ranges, normal and log-normal distribution.

1. INTRODUCTION

In recent development, the importance of bioequivalence cannot be discounted. Because in pharmaceutical industries, regulatory agencies permit manufacturer to market a generic drug if they can prove that the generic product is bioequivalent to the brand-name product, the assumption is that the bioequivalent drug will produce the same therapeutic effect as the brand-name product. Furthermore, no need to perform clinical trial to demonstrate the safety and efficacy of the generic product when the generic product manufacturer has already been proved it bioequivalent to the brand-name product [1].

Bioequivalence study refers to the comparison of pharmacokinetic parameters such as AUC, Cmax etc., as observed in the experiments related to the formulations to be tested. Two formulations are claimed as bioequivalent when their amount and speed of absorption do not exhibit statistically significant difference when administrated at the same molar dose of active ingredient, under similar experimental conditions [2].

According to several regulatory recommendations Chow and Liu [3] have defined different average bioequivalence criteria, discussed as follows.

75/75 Rule

Bioequivalence is concluded when at least 75% of subject ratios (relative individual bioavailability of the test (T) formulation to the reference (R) formulation) fall within range 75%----125%.

80/20 Rule

Bioequivalence is claimed when test and reference formulation means are not statistically significant different and there is at least 80% power for the detection of 20% difference of the reference average.

±20% Rule

The test formulation is claimed as bioequivalent to the reference when the average bioavailability of the test formulation is within 20% of that of the reference formulation with a certain assurance.

80%/125% Rule

Regulatory agencies (e.g., FDA [4] and FDA [5]) recommend the use of logarithmic transformation for AUC and Cmax for the bioequivalence analysis. According to 80% / 125% rule, the test and reference formulations are claimed as bioequivalent to each other when the average bioavailability of the test formulation is within (80%, 125%) that of the reference formulation. This 80%/125% rule which is based on the ratios of two average formulations also enables bioequivalence to be explained in terms of a difference rather than ratio. When the linear (original) scale is considered this 80%/125% rule is not symmetric about 1 where the maximum probability of concluding average bioequivalence occurs. However, on the logarithmic scale, this criterion has a range of -0.2231 to 0.2231 and it is symmetric about 0 where the probability of concluding average bioequivalence is at maximum. On logarithmic scale two formulations are considered to be bioequivalent if 90% confidence interval for the ratio of two formulations’ average are within the range of (80% ---125%) or 90% confidence interval for the difference of two formulations’ average are within the range of (-0.2231----0.2231).
Pharmacokinetic (PK) measures assessed by bioequivalence are derived directly from the drug concentration–time curve, which is described by the quantification of a given number of biological samples related to previously established collection times.

The most important PK measure assessed is the area under the drug concentration–time curve (AUC), frequently used to measure the total amount of drug absorbed by the body after a drug administration. Maximum concentration ($C_{\text{max}}$) is the measure that represents the major drug concentration observed and is directly proportional to the total drug amount absorbed by the body. Another PK measure is $T_{\text{max}}$ that is the collection time in which the maximum concentration was observed and is related to the drug absorption speed [6, 7].

Commonly used design for the evaluation of bioequivalence of two formulations (T=test and R=reference) is two-period, two-sequence (2X2) crossover design [8, 9]. In 2x2 crossover design, each study subject is administrated a formulation of a drug and then switched over to other formulation after a washout period of time [3]. An adequate washout period following the administration of any period ascertains that drug given is completely eliminated before administration of the next drug. Therefore, in 2X2 crossover design, each study subject is allocated either TR or RT sequences in two periods at random. More explicitly, subjects allocated to RT (TR) sequence receive formulation R (T) in the first period and formulation T (R) in the second period [10].

2. MULTIPlicative MODEL

The distribution of PK measures e.g., AUC and $C_{\text{max}}$ are often positively skewed instead of normal, that lead to the violation of homogeneity of variances assumption. In such situation, the assessment of average bioequivalence based upon raw data model may not be suitable. Therefore, the logarithmic transformation is usually considered in order to achieve the normality of data and additive model with relatively homogeneous variances [3]. The multiplicative model (or log-transformed model) is given as:

$$X_{ik} = \ln(X_{Tik}) = \ln(\mu) + S_{ik} + P_j + F_{(i,j,k)} + e_{ijk}$$  \hspace{1cm} (1)

$$Y_{ik} = \ln(\frac{X_{Tik}}{X_{Rik}}) = \ln(\frac{\mu}{\mu} + S_{ik} + P_j + F_{(i,j,k)} + e_{ijk})$$  \hspace{1cm} (2)

Where $\mu$ is the overall mean, $P_j$ is the fixed effect of the $j^{th}$ period with $\sum P_j = 0$, $F_{(i,j,k)}$ is the fixed effect of the formulation in the $k^{th}$ sequence at $j^{th}$ period with $\sum F_{(i,j,k)} = 0$. $C_{(j-1,k)}$ is the fixed carryover effect observed for the formulation in $k^{th}$ sequence at $(j-1)^{th}$ period with $\sum C_{(j-1,k)} = 0$, $S_{ik}$ is the random effect for the $i^{th}$ subject in $k^{th}$ sequence, and $e_{ijk}$ is the random error. From the above multiplicative model it can be seen that $\ln(X_{Tik}) = \ln(\mu) + S_{ik} + P_j + F_{(i,j,k)} + e_{ijk}$. If it is assumed that $\{S_{ik}\}$ and $\{e_{ijk}\}$ are independently and normally distributed with covariance structure defined by Chow and Liu [3] then $X_{Tik}$ follows a lognormal linear model [11, 12]. Equation (1) can be expressed as

$$X_{Tik} = \exp(\mu + S_{ik} + P_j + F_{(i,j,k)} + C_{(j-1,k)} + e_{ijk})$$

The lognormal linear model in (1) requires two assumptions, (i) the log transformed data can be expressed by an additive model and, (ii) the distributions of $\{S_{ik}\}$ and $\{e_{ijk}\}$ are independent and normal. Therefore, it is reasonably substantial to verify these assumptions on the log-transformed and the linear scale [13].

Usually the ratio of means of PK parameter between test and reference formulations is considered as the measure of average bioequivalence. As mentioned above, the situation where the distribution of PK parameter is not normal, some researchers suggest that the ratio of medians (i.e., $M(X_T)/M(X_R)$), instead of the ratio of means, should be used as an alternative measure of average bioequivalence under the fact that in such situations the medians are better representative of central location [3].

3. USE OF GEOMETRIC MEAN

When the logarithmic scale is used, the FDA [4] requests that in the assessment of average bioequivalence the ratio of geometric mean of two formulations on the linear scale should be considered [16].

This use of geometric means is the fact that, geometric mean is an approximately unbiased estimator of median rather than mean.

A positive-valued random variable $X$ follow lognormal distribution when logarithm of $X$ is normally distributed

$$Y = \ln(X) \sim N(\mu, \sigma^2)$$

$$\ln(X_T / X_R) = \ln(X_T) - \ln(X_R) \sim (\mu_T - \mu_R, \sigma^2_T + \sigma^2_R - 2\sigma_{\text{int}}) \hspace{1cm} (3)$$

$$\ln(Y_T / Y_R) = \ln(Y_T) - \ln(Y_R) \sim (\mu_T - \mu_R, \sigma^2_T + \sigma^2_R - 2\sigma_{\text{int}})$$
A random sample $X_1, X_2, \ldots, X_n$ is drawn from lognormal distribution with corresponding mean $\mu$ and variance $\sigma^2$ on the log transformed scale, i.e., $Y_i = \ln(X_i) \sim N(\mu, \sigma^2)$, the arithmetic mean $\bar{Y}$ is the point estimate of $\mu$.

**Theorem 1:** An Exponential transformation of the estimate $\bar{Y}$ results in the geometric mean of the untransformed variable, i.e.

$$\exp(\bar{Y}) = \sqrt[n]{\exp(X_i)} = G.M. (X_i)$$

**Proof:**

$$\exp(\bar{Y}) = \exp\left(\frac{\sum X_i}{n}\right) = \prod \exp\left(\frac{X_i}{n}\right) = \prod \exp\left(\ln \frac{X_i}{n}\right)$$

$$= \prod \exp\left(\ln X_i\right) = \prod X_i^{1/n} = \sqrt[n]{\prod X_i}$$

Hence we can write as

$$\exp(\bar{Y}) = G.M. (X_i)$$

**Theorem 2:** Geometric mean is not an unbiased estimator of the population mean on the linear scale.

**Proof:**

Consider, $E\left(\prod X_i\right)$

$$E\left(\prod X_i\right) = E\left(\prod X_i^{1/n}\right) = \prod E\left(X_i^{1/n}\right)$$

Now consider $E\left(X_i^{1/n}\right)$

The density function of log normal random variable $x_i$ is

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \frac{1}{x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) dx$$

It can be easily proved that

$$E(X_i) = \exp\left(\mu + \frac{\sigma^2}{2}\right) and M(X_i) = \exp(\mu)$$

Where $E(X_i)$ and $M(X_i)$ are mean and median of log normal random variable $X_i$

$$E\left(X_i^{1/n}\right) = \frac{1}{\sigma \sqrt{2\pi}} \frac{1}{x^{1/n}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) dx$$

Let $\ln X = t \Rightarrow x = \exp(t)$

$$dx = \exp(t) dt$$

Now

$$E\left(X_i^{1/n}\right) = \int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} \frac{1}{x^{1/n}} \exp\left(-\frac{(t - \mu)^2}{2\sigma^2}\right) \exp\left(-\frac{(t - \mu)^2}{2\sigma^2}\right) dt$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(t - \mu)^2}{2\sigma^2}\right) dt$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(t - \mu - \sigma^2/2)^2}{2\sigma^2}\right) dt$$

$$= \exp\left(\frac{\mu + \sigma^2}{2n}\right) \int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(t - \mu - \sigma^2/2)^2}{2\sigma^2}\right) dt$$

$$= \exp\left(\frac{\mu + \sigma^2}{2n}\right)$$

Since $\int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(t - \mu - \sigma^2/2)^2}{2\sigma^2}\right) dt = 1$ with

mean $= \mu + \frac{\sigma^2}{2}$ and variance $= \sigma^2$

Therefore $E\left(X_i^{1/n}\right) = \exp\left(\frac{\mu + \sigma^2}{2n}\right)$

Now we can write as

$$E\left(\prod X_i\right) = E\left(\prod X_i^{1/n}\right) = \prod E\left(X_i\right) = \prod (\mu + \sigma^2/2n)$$

$$E\left(\prod X_i\right) = E(G.M. (X_i)) = \exp\left(\mu + \frac{\sigma^2}{2n}\right)$$

Hence it is proved that the geometric mean on the linear scale is not an unbiased estimator of the population mean. The associated bias factor can be described as

Using the equation (5)
\[ E(\exp(\bar{Y})) = \exp\left( \mu + \frac{\sigma^2}{2n} \right) \]

\[ E(\exp(\bar{Y})) = \exp\left( \mu + \frac{\sigma^2}{2n} \right) \cdot \exp\left( -\frac{\sigma^2}{2}(1-1/n) \right) \]

\[ E(\exp(\bar{Y})) = E(X_i) \cdot \exp\left( -\frac{\sigma^2}{2}(1-1/n) \right) = (true \ median)(bias \ factor) \]

Since for \( \sigma^2 > 0 \) the bias factor \( \exp\left( -\frac{\sigma^2}{2}(1-1/n) \right) \) is always smaller than 1, therefore,

\[ E(\exp(\bar{Y})) = E(X_i) \cdot \exp\left( -\frac{\sigma^2}{2}(1-1/n) \right) < E(X_i) \]

Geometric mean is also not unbiased estimator of population median and associated biased factor can be described as

\[ E(\exp(\bar{Y})) = \exp(\mu) \exp\left( \frac{\sigma^2}{2n} \right) = (true \ median)(bias \ factor) \]

Gilbert [17] mentioned that this bias factor is positive and as \( n \) increases the bias factor decrease. Hence, for the large sample size \( \frac{1}{n} = 0 \) the geometric mean is approximately unbiased estimator of the population median.

\[ E(\exp(\bar{Y})) = \exp\left( \mu + \frac{\sigma^2}{2n} \right) = \exp(\mu) = M(X) \]

When true population mean is estimated by computing geometric mean, the bias factor is less than 1 and does not go to zero when \( n \) increases. Whereas, when true population median is estimated by computing geometric mean, the bias factor decreases as \( n \) increases, therefore, geometric mean is approximately unbiased estimator of population median. This property of geometric mean became a rationale for using the geometric mean ratio of test and reference formulations for the point estimate in average bioequivalence in the situation where PK parameters are not normally distributed and the ratio of two formulation medians is a better choice to measure the average bioequivalence.

The test formulation is said to be bioequivalent to the reference one if a (1-2\( \alpha \)) \times 100\% confidence interval for the difference or the ratio of formulation means fall between the bioequivalence ranges recommended by regulatory authorities. Usually level of significance considered as 5\%, consequently 90\% confidence interval for the difference or ratio is used for evaluation of average bioequivalence.

**Linear Scale (without Transformation)**

To claim, the test formulation is average bioequivalent to the reference formulation, it is usually required that the confidence interval for the difference of two formulations means be within \( \pm 20\% \) of the reference formulation (or the ratio of two formulations means be within (80\%---120\%)).

i.e., the average bioequivalence interval using the difference of two formulations means is expressed as

\[-0.2 \mu_R < (\mu_T - \mu_R) < 0.2 \mu_R \]

(6)

In common practice, the limits \( \pm 0.2 \mu_R \) shall be estimated as they depend on the population parameter, the usual choice is \( \pm 0.20 \bar{Y}_R \). Where \( \mu_R \) and \( \mu_T \) are reference and test formulation population means on the linear scale respectively and \( \bar{Y}_R \) is the reference formulation sample mean.

Dividing the above equation (6) by \( \mu_R \) give the average bioequivalence interval for the ratio of two formulation means

\[-0.20 < \frac{(\mu_T - \mu_R)}{\mu_R} < 0.20 \]

\[0.80 < \frac{\mu_T}{\mu_R} < 1.20 \]

(7)

**Logarithmic Scale (with Transformation)**

Let \( X \) denote the pharmacokinetic measure on the linear scale, now \( \mu \) denotes the geometric mean of \( X \) and \( \mu \) denotes the arithmetic mean of log normally distributed pharmacokinetic measure (\( y = \ln X \)).

When logarithmic scale is used, the FDA [4] suggests that to conclude average bioequivalence the interval of the ratio of geometric mean (computed on the linear scale) of two formulations be within range (80\%---125\%), i.e.,

\[0.80 < \frac{\mu_T}{\mu_R} < 1.25 \]

(8)

The use of the range 80\% to 125\% is the fact that there is a correspondence to the symmetrical average bioequivalence interval for the difference of means on the logarithmic scale.
After taking the logarithm, the equation (8) becomes

\[ \ln(0.80) < \ln \left( \frac{\mu_T}{\mu_R} \right) < \ln(1.25) \]  

(9)

Using the equation (5)

\[-0.22314 < \ln \left( \frac{e^{\mu_T}}{e^{\mu_R}} \right) < 0.22314\]  

(10)

Now it can be easily written as

\[-0.22314 < (\mu_T - \mu_R) < 0.22314\]  

(11)

Where \( \mu_R \) and \( \mu_T \) are the test and reference formulation population means on the logarithmic scale.

4. EXAMPLE

Hauschke, Steinijans [18] considered the following example of BE study from [19]. This was a two-period, two-sequence crossover study with a one-week washout period, where two capsules each of 200 and 300 mg theophylline used as the reference and two capsules of 500 mg theophylline as test formulation.

Eighteen subjects were randomized equally into two sequences RT and TR. Blood samples were taken prior to drug administration and after 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 28, 36, 40, 44 and 48 hours after the administration. One of the primary pharmacokinetic parameter was AUC (0-inf). The data are given in Table 2.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Difference</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>(-0.2 \mu_x &lt; (\mu_T - \mu_R) &lt; 0.20 \mu_x)</td>
<td>(0.80 &lt; \frac{\mu_T}{\mu_R} &lt; 1.20)</td>
</tr>
<tr>
<td>Logarithmic</td>
<td>(-0.22314 &lt; (\mu_T - \mu_R) &lt; 0.22314)</td>
<td>(0.80 &lt; \frac{\mu_T}{\mu_R} &lt; 1.25)</td>
</tr>
</tbody>
</table>

Table 2: Area under the curve AUC (0-inf) for the test (T) and reference (R) formulations

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Sequence</th>
<th>Period I</th>
<th>Period II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TR</td>
<td>228.04</td>
<td>288.79</td>
</tr>
<tr>
<td>2</td>
<td>RT</td>
<td>339.03</td>
<td>329.76</td>
</tr>
<tr>
<td>3</td>
<td>TR</td>
<td>288.21</td>
<td>343.37</td>
</tr>
<tr>
<td>4</td>
<td>RT</td>
<td>242.64</td>
<td>258.19</td>
</tr>
<tr>
<td>5</td>
<td>RT</td>
<td>249.94</td>
<td>201.56</td>
</tr>
<tr>
<td>6</td>
<td>TR</td>
<td>217.97</td>
<td>225.77</td>
</tr>
<tr>
<td>7</td>
<td>TR</td>
<td>133.13</td>
<td>235.89</td>
</tr>
<tr>
<td>8</td>
<td>RT</td>
<td>184.32</td>
<td>249.64</td>
</tr>
<tr>
<td>9</td>
<td>TR</td>
<td>213.78</td>
<td>215.14</td>
</tr>
<tr>
<td>10</td>
<td>TR</td>
<td>248.98</td>
<td>245.48</td>
</tr>
<tr>
<td>11</td>
<td>TR</td>
<td>163.93</td>
<td>134.89</td>
</tr>
<tr>
<td>12</td>
<td>RT</td>
<td>209.3</td>
<td>231.98</td>
</tr>
<tr>
<td>13</td>
<td>RT</td>
<td>207.4</td>
<td>234.19</td>
</tr>
<tr>
<td>14</td>
<td>TR</td>
<td>245.92</td>
<td>223.39</td>
</tr>
<tr>
<td>15</td>
<td>RT</td>
<td>239.84</td>
<td>241.25</td>
</tr>
<tr>
<td>16</td>
<td>RT</td>
<td>211.24</td>
<td>255.6</td>
</tr>
<tr>
<td>17</td>
<td>TR</td>
<td>188.05</td>
<td>169.7</td>
</tr>
<tr>
<td>18</td>
<td>RT</td>
<td>230.36</td>
<td>256.55</td>
</tr>
</tbody>
</table>
Descriptive statistics such as mean and standard deviations for test and reference formulations in both periods I and II are presented in Table 3.

Point estimates and 90% confidence intervals for difference and ratio of formulation means on linear and logarithmic scales are presented in Tables 4 and 5 along with bioequivalence decisions.

It is obvious from Tables 4 and 5 that the confidence intervals for the difference and ratio of two formulations means are within the bioequivalence ranges described above, therefore, it is concluded that reference formulation is the bioequivalent to the test formulation.

5. DISCUSSION

Under the assumptions of a lognormal linear model, the exponentiation of the confidence interval for the difference of test and reference formulation means based on the log-transformed scale gives the confidence interval for the ratio of geometric means of two formulations on the linear scale. As recommended by the FDA, for the assessment of average bioequivalence, the ratio of geometric means must be used when linear data is apart from normality assumption. Therefore, one must have a clear concept about the assessment of bioequivalence on the logarithmic scale. In this research, we tried to describe the rationale, why the ratio of geometric means of test and reference formulations is used when data exhibit violation of normality assumption. We also tried to clarify the concept of average bioequivalence intervals based on the difference or ratio of test and reference formulation means.

REFERENCES


