Research Paper

Volume : 4 | Issue : 7 | July 2015 • ISSN No 2277 - 8179

Statistical Models in Replicated Crossover Design in Bioequivalence Clinical Trial: Fixed Effect Model Versus Random Effect Model

Statistics

KEYWORDS : Bioequivalence, Replicated crossover design, ANOVA

Sujata Suvarnapathki Department of Statistics, Ramnarain Ruia College, Matunga, Mumbai

Manik Khaparde Department of Statistics PGTD, Rashtrasanta Tukdoji Maharaj Nagpur University, Nagpur

ABSTRACT

In Analysis of Variance, there are two types of factors fixed effect and random effect. An appropriate type of effect is chosen depending on the context of the problem.

In Fixed effect modelling, the interest lies in comparison of the specific levels e.g. Formulation or treatment for a particular drug product.

In random effect modelling, the levels of the factor are drawn randomly from universe.

In case of Bioequivalence studies, EMEA and US FDA differ on the view of treating subject effect. EMEA assumes subject effect as fixed whereas US FDA assumes it to be random.

In this paper, three methods for statistical analysis of Pharmacokinetic parameters are compared for random effect and fixed effect model for replicated crossover design in a bioequivalence study.

Conclusions: The results of the statistical analyses indicate that 90% confidence interval for the three approaches are not same. In general, Method C gives wider confidence intervals.

1. Introduction:

Fixed effects can be thought of as treatment levels that we have selected for inclusion in a study, (in Bioequivalence, its R i.e. Reference and T i.e. Test) which are the only levels of the variable in question in which we have an interest. In an experiment, we might have a treatment group and a control group.

Random effect is an effect where there is source of random variation or experimental units e.g., individuals drawn at random from a population for a clinical trial. Random effects estimate the variability. A random effect(s) model, also called a variance components model, is a kind of hierarchical linear model.

Bioavailability is a measurement of the extent of a therapeutically active medicine that reaches the systemic circulation and is therefore available at the site of action.

The average bioequivalence criterion stipulates that two drugs are to be considered bioequivalent when the 90% confidence interval, considering the average bioavailability of the test drug (T) and the reference drug (R) and the T/R ratio, is between 80% and 125%, for data converted to the logarithmic scale.

The two drugs are bioequivalent if the 90% Confidence Interval for the ratio of geometric means of test preparation to standard preparation (T/R ratio) lies in 80% to 125%.

A two-sequence (RTRT, TRTR), four-period crossover design is used for the assessment of bioequivalence between reference formulation (R) and a test formulation (T) of a drug product. Such a design is referred as Replicated Crossover Design. In this design, each subject is randomly assigned to either of two sequences at four dosing periods. The dosing periods are separated by washout period of sufficient length.

In this paper different approaches for statistical analysis of replicated crossover bioequivalence study are compared.

Three commonly suggested approaches are used to carry out the statistical analysis of such design and are compared. Replicated cross over design is most acceptable design in Bioequivalence studies for highly variable drug products.

In Method A, subject effect is considered to be fixed. In Method B, subject effect is considered to be random. Method C also considers subject effect to be random rather than fixed but also allows a different subject effect for each formulation.

The approach of Method A is feasible even for unbalanced replicate design studies. The advantage of this approach is that it is straightforward and that it appears to be software and software option independent.

A simple linear mixed model, which assumes identical within-subject variability (Method B), may be acceptable as long as results obtained with the two methods do not lead to different regulatory decisions. However, in borderline cases and when there are many included subjects who only provide data for a subset of the treatment periods, additional analysis using method A might be required.

Method C is such that the model allows a different subject effect for each formulation (i.e. a subject by formulation interaction), and therefore has 5 variance terms (within subject for reference, within subject for test, between subject for test, between subject for reference, covariance for between subject test and reference – the last three are combined to give the subject ×formulation interaction variance component.)

For highly-variable drugs it is recommended to estimate the within subject variance using data from the reference formulation only.

2. Methodology:

Three data sets are simulated assuming Multivariate Log Normal distribution for pharmacokinetic parameters Cmax. Literature data is used to estimate the geometric mean ratio. Intra subject coefficient of variation is varied as 0.30, 0.35, and 0.40. The estimated sample sizes are 18, 24, and 32.

2.1 Method A: In this method Subject effect is assumed to be fixed effect.

Model for replicated crossover design can be rewritten as follows

$$Y_{ijl} = \mu + S_i + F_j + P_l + e_{ijl} \qquad j,l = 1,2,3.....f, \quad i = 1,2.....N$$

Where Y_{ijl} is the response variable on the i^{th} subject in the l^{th} period under the j^{th} Formulation. μ is the overall mean.

Fj is the fixed effect of the j^{th} formulation with $\sum_j F_j = 0$

P_l is the fixed effect of the l^{th} period with $\sum_l P_l = 0$

Si is the fixed effect of the i^{th} subject, e_{ijl} is the error term.

σ_e^2 is the variance component related to subject effect (Within subject error variance).

For this model there is only one variance term estimated, σ_w^2 the within subject variability.

2.2 Method B: In this method Subject effect is assumed to be random effect.

This means there are two variance terms estimated σ_w^2 and σ_b^2, as a distribution is also fitted to the between subject variability. If subject is a fixed effect (as in the previous model) each subject is treated as being selected in some way rather than being sampled from a random distribution and a subject effect is estimated individually for each patient as is done for the period effect.

This model will give the same results as Method A if all subjects included in the analysis provide data for all treatment periods.

Model for replicated crossover design can be rewritten as follows.

$$Y_{ijl} = \mu + S_i + F_j + P_l + e_{ijl} \qquad j,l = 1,2,3.....f, \;\; i = 1,2.....N$$

Where Y_{ijl} is the response variable on the i^{th} subject in the l^{th} period under the j^{th}

Formulation.

μ is the overall mean.

Fj is the fixed effect of the j^{th} formulation with $\sum_j F_j = 0$

P_l is the fixed effect of the l^{th} period with $\sum_l P_l = 0$

Si is the random effect of the i^{th} subject, e_{ijl} is the error term.

σ_s^2 is the variance component related to subject effect (Between subject error variance).

σ_e^2 is the variance component related to subject effect (Within subject error variance).

2.3 Method C: In this method Subject effect is assumed to be random effect.

Model for replicated crossover design can be rewritten as follows.

$$Y_{ijl} = \mu + S_i + F_j + P_l + e_{ijl} \qquad j,l = 1,2,3.....f, \;\; i = 1,2.....N$$

Where Y_{ijl} is the response variable on the i^{th} subject in the l^{th} period under the j^{th} Formulation. μ is the overall mean.

Fj is the fixed effect of the j^{th} formulation with $\sum_j F_j = 0$

P_l is the fixed effect of the l^{th} period with $\sum_l P_l = 0$

Si is the random effect of the i^{th} subject, e_{ijl} is the error term.

This model allows a different subject effect for each formulation (i.e. a subject by formulation interaction), and therefore has 5 variance terms (within subject for reference, within subject for test, between subject for test, between subject for reference, covariance for between subject test and reference – the last three are combined to give the subject ×formulation interaction variance component.)

3. Numerical Analysis:

SAS codes are used to analyse the data sets with sample size 18, 24 and 32 that are generated using multivariate normal distribution.

For Method A, PROC GLM procedure in SAS is used and for Method B and C, PROC MIXED procedure in SAS is used.

The comparison of three methods is based on ratio of least square means (T/R) and its 90% confidence interval.

The results are as follows.

Table 3.1: Table for Original and Modified Confidence Interval (With missing observations) for sample size N=18.

Method	N	90% Confidence Interval		90% Confidence Interval(With Missing data)	
		Lower Limit	Upper Limit	Lower Limit	Upper Limit
A	18	91.53	102.47	91.39	102.31
B	18	91.53	102.47	91.39	102.27
C	18	89.99	104.23	89.50	104.35

Table 3.2: Table for Original and Modified Confidence Interval (With missing observations) for sample size N=24.

Method	N	90% Confidence Interval		90% Confidence Interval(With Missing data)	
		Lower Limit	Upper Limit	Lower Limit	Upper Limit
A	24	97.64	108.79	97.32	108.76
B	24	97.64	108.79	97.24	108.66
C	24	97.42	109.03	96.88	108.82

Table 3.3: Table for Original and Modified Confidence Interval (With missing observations) for sample size N=32.

Method	N	90% Confidence Interval		90% Confidence Interval(With Missing data)	
		Lower Limit	Upper Limit	Lower Limit	Upper Limit
A	32	94.38	104.48	94.15	104.42
B	32	94.38	104.48	94.03	104.29
C	32	97.42	109.03	96.88	108.82

3.4 Discussion:

The Guideline on the Investigation of Bioequivalence CPMP/QWP/EWP/1401/98 Rev. 1) recommends analyzing bioequivalence studies using ANOVA and specifying all factors, including subjects, as fixed rather than random. The analysis presented above show that this approach (Method A) is feasible even for unbalanced replicate design studies. The advantage of this approach is that it is straightforward and that it appears to be statistical software and software option independent. A simple linear mixed model, which assumes identical within-subject variability (Method B), may be acceptable as long as results obtained with the two methods do not lead to different regulatory decisions. However, in borderline cases and then there are many included subjects who only provide data for a subset of the treatment periods, additional analysis using method A might be required.

For highly-variable drugs it is recommended to estimate the within subject variance using data from the reference formulation only.

When there are no subjects with missing treatment periods, the results from methods A and B are identical and the point estimate is the same for all three approaches. Method C gives wider intervals.

The results are generally very similar although missing treatment periods for some causes the results to be different for all three approaches.

Volume : 4 | Issue : 7 | July 2015 • ISSN No 2277 - 8179

REFERENCE

[1] EMEA Guidelines for establishing Bioequivalence | [2] Guidelines & Guidance Documents | Guidance Document: Conduct and Analysis of Comparative Bioavailability Studies | [3] Sanford Bolton and Charles Bon: Pharmaceutical Statistics: Practical and Clinical applications | [4] Shein-Chung Chow and Jen-Pei Liu: Design and Analysis of Bioavailability and Bioequivalence Studies | [5] US FDA guidelines for establishing | Bioequivalence