Resolution - RE n. 898, of May 29, 2003

D.O.U. 02/06/2003

The Deputy of the Collegiate Board of Directors of the Brazilian Sanitary Surveillance Agency, in the use of the attribution vested in him by Administrative Order n. 238, of March 31, 2003,

WHEREAS

provided in Article 111, clause II, item “a” of paragraph 3 of the Bylaws approved by Administrative Order 593, of August 25, 2000, re-published in the Federal Official Journal of December 22, 2000

that the matter was submitted to the examination of the Collegiate Board of Directors, which approved the matter in a meeting held on March 6, 2003, decides:

Article 1 - To determine the publication of the " Guide for planning and carrying out the statistical stage of relative bioavailability/ bioequivalence studies”, attached.

Article 2 – Resolution RE n. 484, of March 19, 2002, is hereby revoked.

Article 3 - This Resolution enters into force on the date of its publication.

DAVI RUMEL

ANNEX

GUIDE FOR PLANNING AND CARRYING OUT THE STATISTICAL STAGE OF STUDIES OF RELATIVE BIOAVAILABILITY/BIOEQUIVALENCE

1. Introduction

The aim of this guide is to provide general recommendations regarding statistical analysis in relative bioavailability/bioequivalence studies.

2. Planning

One of the criteria for choosing an appropriate design is to verify whether the selected design can identify and isolate the inter-individual variability in the data analysis. Any design that removes this variation of the comparison between formulations is appropriate.

The experimental planning most often used in relative bioavailability/bioequivalence assays is the crossover design, whose details shall be discussed in this guide.

2.1 Washout and carry-over effects

It is important to introduce the concepts of washout and carry-over effects in a crossover design , since the presence of carry-over effects has a great impact on the statistical inference of bioequivalence between formulations.

Washout is defined as a time interval long enough between two periods of administration for the carry-over effect of a formulation administered in a period to be eliminated until the next one.

A crossover design should be used when there is no carry-over effect in the treatments. If a drug has a long half-life or if the washout interval between periods of treatment is very short, the effect of the drug may continue after the end of the dosing period. In this case, it is necessary to distinguish the effect of the drug from the carry-over effect The effect of the drug is the one observed during the period in which it is administered.

2.2 Description of the planning

A crossover design is a modified randomized block design in which each block receives more than one formulation of the same drug in different periods. A block can be an individual or a group of individuals. The individuals in each block receive a different formulation sequence. The advantages in using this design for relative bioavailability/bioequivalence studies are:

· each individual is his/her own control, which allows comparison of the individual with him/herself, for the different formulations;
· the inter-individual variability is removed from the comparison between formulations, which makes the test of treatment difference generally more powerful;
· with appropriate randomization of individuals for the administration sequence of the formulations, the design produces the best estimates unbiased by the difference (or ratios) between formulations.

2.3 Considerations regarding basic design

It is recommended that a basic design for an *in vivo* bioavailability study consider:

· scientific questions to be answered;
· nature of the reference material and the pharmaceutical form to be tested;
· availability of analytical methods;
· considerations regarding the benefit of the test in human beings.

Moreover, some specific considerations for a relative bioavailability/bioequivalence study are provided below.

2.3.1. Experiment design

For a relative bioavailability/bioequivalence study (simple or multiple dose) a crossover design should be adopted, unless a parallel design or some other is more appropriate for valid scientific reasons. In the case of parallel design, each individual receives at random only one of the formulations.

Adequate design experiment should be aimed at minimizing the variability that can result from several sources:

· inter-individual variability.
· intra-individual variability.
· effect of the periods, which can be caused by carry-over action of previous treatments;
· experiment error.
·variability associated to different treatments, such as administration of different products or dosages.

2.3.2. Randomization

Valid statistical inferences are usually based on the assumptions that the errors of the model employed are independently distributed random variables, which can be ensured through the randomization. The randomization form is carried out according to the design to be used in the study.

2.3.3. Sampling schedule

2.3.4. Washout

2.3.5. Number of subjects

The number of healthy subjects must always ensure enough statistical power to guarantee the reliability of the results of the relative bioavailability/bioequivalence study.

2.4 Types of design

THIS SECTION DESCRIBES DESIGNS COMMONLY USED IN RELATIVE BIOAVAILABILITY/BIOEQUIVALENCE STUDIES.

2.4.1. Crossover designs employing two drug products (T = test; R = reference)

a) 2X2 crossover design

This is a conventional non-replicated design with two formulations, two periods and two sequences, which may be represented as follows:

	Período	
Seqüência	1	2
1	R	T
2	T	R

Each individual randomly placed in RT or TR sequence in two periods. That is, individuals placed in RT (TR) sequence receive formulation R (T) in the first administration period and formulation T (R) in the second. The periods are separated for adequate washout.

Randomization for a 2x2 crossover study can be done through tables of random numbers or randomization procedures implemented through statistics software.

b) Replicated crossover design

This design is recommended for relative bioavailability/bioequivalence studies of products with high variability drugs (intra-individual variation coefficient ≥ 30%), including those that are immediate release, modified release and other oral administration products.

For these designs, the same batches of the test and reference formulations must be used for replicated administration. The periods must be sufficiently spaced to guarantee no carry-over effect.

The replicated crossover designs most commonly used to compare two formulations are:

I. Design with four sequences and two periods (Bataam design):

	Periodo	
Seqüência	1	2
1	T	T
2	R	R
3	R	T
4	T	R

II Design with two sequences and four periods:

	Período			
Seqüência	1	2	3	4
1	T	R	R	T
2	R	T	T	R

III Design with four sequences and four periods:

	Período			
Seqüência	1	2	3	4
1	T	T	R	R
2	R	R	T	T
3	T	R	R	T
4	R	T	T	R

IV Design with two sequences and three periods:

	Período		
Seqüência	1	2	3
1	T	R	T
2	R	T	R

Or

	Período		
Seqüência	1	2	3
1	T	R	R
2	R	T	T

A higher number of subjects is recommended for the three period design, compared to the four period design, in order to reach the same statistical power for the test.

c) Crossover design employing three drug products (Williams design, using T1 = test 1; T2 = test 2; R = reference)

In order to compare three formulations of a drug, there are three possible pairs of comparisons: formulation 1 versus formulation 2, formulation 1 versus 3 formulation and formulation 2 versus formulation 3. When the number of formulations to be compared is large, more sequences and

consequently more individuals are necessary, which can be impracticable. A practical design proposed by Williams (1949) has balancing properties and requires few sequences and periods. A design is said to be balanced if it meets the following conditions:

· each drug product is applied only once for each subject;
· in each period, the number of subjects receiving each drug product must be the same;
· the number of subjects receiving drug product i in some period followed by drug product j in the following period is the same for all $i \neq j$.

A Williams design is illustrated as follows:

	Periodo		
Seqüência	1	2	3
1	R	T2	T1
2	T1	R	T2
3	T2	T1	R
4	T1	T2	R
5	T2	R	T1
6	R	T1	T2

d) Crossover design for four drug products (Williams design):

	Periodo			
Seqüência	1	2	3	4
1	R	T3	T1	T2
2	T1	R	T2	T3
3	T2	T1	T3	R
4	T3	T2	R	T1

2.5 Selection of experiment design

It is important to selecting an appropriate design when planning a relative bioavailability/bioequivalence study. This depends on several factors, such as:

· number of formulations to be compared;
· characteristics of the drug and its bioavailability;
· objective of the study;
· inter and intra individual variability;
· duration of the study and number of periods employed;
· cost of adding a subject compared to cost of adding a period;
· dropout rate.

The analysis of the data, the interpretation of the results and the determination of bioequivalence between the formulations, depend directly on the selected design. Therefore, all the factors mentioned above must be carefully assessed for an appropriate design to be chosen.

3 Statistical Analysis

3.1 Logarithmic transformation

3.1.1 General procedure

This guide recommends that the values of the parameters (ASC and Cmax) be transformed using natural logarithms or common logarithms in 10 base. The choice of natural or common logarithms must be consistent and must be specified in the study report.

The limitation of the sample size used in a typical relative bioavailability/bioequivalence study prevents reliable determination of distribution of the set of data. It is not recommended to test normality of distribution of errors after logarithmic transformation, nor should normality of distribution of errors be used as a reason to carry out statistical analysis in the original scales. Justifications must be presented when it is considered best to carry out the statistical analysis in the original scales rather than in the logarithmic scales.

3.1.2 Justifications for use of logarithmic transformation

a) Justification regarding treatment of data

In general, an interesting preliminary comparison in a relative bioavailability/bioequivalence study is the use of the ratio instead of the difference, between the averages of the pharmacokinetic parameters (ASC and Cmax) of the data of the test and reference product. Using logarithmic transformation, the generalized linear model used in the data analysis allows statistical inferences on the difference between the two averages in the logarithmic scale, which can be re-transformed into statistical inferences over the ratio of the two averages in the original scale (Schuirmann, 1989).

b) Justification regarding pharmacokinetics
Westlake (1973, 1988) observed that a multiplicative model is appropriate for pharmacokinetics measurements (ASC and Cmax) in a relative bioavailability/bioequivalence study. Assuming that elimination of the drug is of first order and only occurs from the central compartment, the following equation is obtained after extravascular administration (oral):

$$ASC_{0-\infty} = F.D/CL = F.D/(Vd.K_e),$$

where: F is the absorbed fraction, D is the administered dose, and F.D is the amount of the absorbed drug. CL is the clearance of voluntary data, which is the product of the volume of apparent distribution (Vd) and of the elimination speed constant (Ke). Therefore, the use of ASC as a measurement of the amount of drug product absorbed involves a multiplicative term (CL), which can be considered as a function of the subject. Therefore, Westlake shows that the subject effect is not additive if the data is analyzed in the original scale.

The logarithmic transformation of the ASC results in an additive treatment:

$$\log ASC_{0-\infty} = \log F + \log D - \log V - \log Ke.$$

Similar arguments have been made for Cmax.

3.2 Data analysis

The parametric methods of generalized linear models are recommended for the analysis of pharmacokinetics measurements transformed into logarithms in a relative bioavailability/bioequivalence study. A variance analysis (ANOVA) must be used in the pharmacokinetic parameters ASC and Cmax using generalized linear models. Appropriate statistical models according to the design chosen in the study must be employed. For example, in a conventional study of the 2x2 crossover type, the statistical model usually includes sequence factors, subject inside sequence, period and treatment. The result should be represented as follows (ANOVA table):

Source of variation	Degree of freedom	Mean squares	Statistic F	P-Value
Sequence	1	(1)	Fr=(1)/(2)	
Subject (sequence)	N-2	(2)		
Period	1	(3)	Fp=(3)/(5)	
Treatment	1	(4)	Ft=(4)/(5)	
Residuals	N-2	(5)		

The sequence, period and treatment effects must be tested using statistics Fr, Fp and Ft indicated in table ANOVA, respectively. It should be noted that the equality between treatments (inexistence of treatment effect) does not mean bioequivalence between formulations. The construction of the 90% confidence interval for the difference means must be based on the least squares means of the data transformed into logarithmic and the mean squared error of this ANOVA. The antilogarithms of the reliability limits obtained constitute the 90% confidence interval for the ratio of the geometric means between the test and reference products. The average bioequivalence conclusion is reached when this confidence interval is between 80 and 125%. This method is equivalent to the two one-sided tests procedure with null hypothesis of bioinequivalence, with a 5% level of significance.

4. Sequence effect

The presence of sequential (carry-over) effect in the study must be justified. For a 2x2 crossover study, the presence of sequential effect may be accepted if some criteria are observed:

I) it is a single dose study;
II) the study it involves only healthy subjects;
III) the drug is not an endogenous substance;
IV) an adequate washout period was established and the pre-dosage samples do not present any level of detectable drug in all the subjects;
V) the study meets all the scientific and statistical criteria (e.g. protocol, validation, concentration data, statistical analysis, confidence interval).

Under other circumstances, the study it must be redone.

5. Considerations regarding outliers

In the relative bioavailability/bioequivalence study with crossover design, the discrepant points are defined as those where some subjects (outliers) differ notably from the other subjects of the study when comparing test and reference product in the subject himself. The existence of an outlier without violation of the protocol may indicate one of the following situations:

A) failure of the product: in this case, an abnormal response may be present both for the test

product and the reference product;

B) subpopulation: this may occur when an individual represents a population, in which the bioavailability of two products is notably different from the majority of the population.

Due to these facts, in general, the exclusion of outliers is not recommended, mainly for designs that are not replicated.

6. Test power and sample size

The power of the test in a relative bioavailability/bioequivalence study is defined as the probability of accepting the bioequivalence between test and reference product correctly. During the planning stage, one of the most important questions is how many subjects are needed to obtain desired power (for example, 80%) establishing bioequivalence between two formulations within the clinically significant limits (e.g. 20% of the average of the reference). To answer this question, the methodology commonly used is to choose an appropriate sample size through calculation of the power function of the test based on a coefficient estimate of intra-individual variation obtained through literature or a pilot-study.

In literature, there are several methods for determining sample size. In this guide, an approximate formula is presented (Chow & Liu) to calculate sample size of a 2x2 crossover design based on the power function of test per Schuirmann's two one-sided tests procedure for interval hypotheses. The determination of sample size for other types of design must be done analogously.

Let $\theta = \mu_T - \mu_R$, that is, θ measurements the true difference between the averages of the test and reference products. In an average bioequivalence study, considering the 20% rule with $\Delta=0.2\ \mu_R$, to reach a power of (1-β) with α level of significance, the size of the sample for each sequence is:

a) in the case of $\theta = 0$,

$n \geq [t(\alpha, 2n-2) + t(\beta/2, 2n-2)]^2\ (CV/20)^2$;

b) in the case of $\theta \neq 0$,

$n \geq [t(\alpha, 2n-2) + t(\beta, 2n-2)]^2\ [CV/(20-\eta)]^2$,

where $\eta = 100 \times \theta/\mu_R = 100 \times (\mu_T - \mu_R)/\ \mu_R$.

In the two formulas presented above, CV represent the coefficient of intra-individual variation and t(a, b) represents the critical value of distribution t of Student, at the a level of significance with b degrees of freedom.

The total number of subjects needed in a 2x2 crossover design is:

$N = 2n$

Since the degree of freedom (2n-2) presented in the formula is unknown, an iterative procedure is necessary to obtain the value of n. To illustrate this procedure, the following example is presented.

Example: To conduct an average bioequivalence study using a 2x2 crossover design and the 20% rule of difference between two formulations, determine the number of subjects needed to obtain an 80% power detecting a difference of 20% between two formulations. Assuming that the CV in this example is 20%.

In the first place, consider the case where $\theta = 0$,

I) starting with an initial guess: n=12;
II) then, we have the degree of freedom 2n-2=22;
III) using α= 0.05 e β= 0.2, we have

$$t(0.05, 22) = 1.717 \text{ e } t(0.1, 22) = 1.321;$$

IV) $n \geq (1.717 + 1.321)^2 (20/20)^2 \approx 9.2$;
V) now use n = 10 as an initial value for next iteration;
VI) 2n-2 = 18, t(0.05, 18) = 1.734 and t (0.10, 18) = 1.330;
VII) $n \geq (1.734 + 1.330)^2 (20/20)^2 \approx 9.4$;
VIII) as these two iterations resulted in a similar response from 10 subjects for each sequence, a total of 20 subjects is necessary in order to obtain80% power to detect a 20% difference between two formulations for the case of $\theta = 0$.

Now consider the case of $\theta = 0.05\mu_R$,

I) starting with an initial guess: n=14;
II) then, we have the degree of freedom 2n-2=26;
III) using α = 0.05 e β= 0.2, we have

$$t(0.05, 26) = 1.706 \text{ e } t(0.2, 26) = 0.856;$$

IV) $n \geq (1.706 + 0.856)^2 [20/(20-5)]^2 \approx 11.66$;
V) for next iteration, using n = 12 as an initial value;
VI) 2n-2 = 22, t(0.05, 22) = 1.717 and t (0.20, 22) = 0.858;
VII) $n \geq (1.717 + 0.858)^2 [20/(20-5)]^2 \approx 11.79$;
VIII) therefore, a total of 24 subjects is necessary in order to obtain 80% power to detect a 20% difference between two formulations for the case of $\theta = 0.05\mu_R$.

The table below presents the total sample size needed to reach desired power for a 2x2 crossover design of various combinations between θ and CV.

		$100 X (\mu_r - \mu_R) / \mu_R$			
Poder	CV(%)	0%	5%	10%	15%
80%	10	8	8	16	52
	12	8	10	20	74
	14	10	14	26	100
	16	14	16	34	126
	18	16	20	42	162
	20	20	24	52	200
	22	24	28	62	242
	24	28	34	74	288
	26	32	40	86	336
	28	36	46	100	390
	30	40	52	114	448
	32	46	58	128	508
	34	52	66	146	574
	36	58	74	162	644
	38	64	82	180	716
	40	70	90	200	794
90%	10	10	10	20	70
	12	10	14	28	100
	14	14	18	36	136
	16	16	22	46	178
	18	20	28	58	224
	20	24	32	70	276
	22	28	40	86	334
	24	34	46	100	396
	26	40	54	118	466
	28	44	62	136	540
	30	52	70	156	618
	32	58	80	178	704
	34	66	90	200	794
	36	72	100	224	890
	38	80	112	250	992
	40	90	124	276	1098

7. Other considerations

The average bioequivalence criterion is recommended for a comparison between the pharmacokinetics measurement of interest in the majority of relative bioavailability/bioequivalence studies. However, in literature, there are the criteria of individual and population bioequivalence, which can also be very useful in some circumstances.

The average bioequivalence focuses only on the comparison of the population averages of pharmacokinetics measurements of interest and not on the variances of these measurements. This method does not take into consideration the variance associated to the interaction between individuals and formulations, that is, the variation between the averages of the test and reference products due to differences between the individuals. On the other hand, the criteria of individual and population bioequivalence include comparisons beyond the averages, the respective variances associated to the pharmacokinetics measurements of study. The criterion of population bioequivalence evaluates the total variability of the measurements of interest. The criterion of individual bioequivalence includes the intra-individual variability of the test and reference products, as well as the interactions between individuals and formulations.

Hauck & Anderson (1992) present considerations and comparisons of the three types of bioequivalence, as well as indications for the construction of the reliability intervals.

8. Bibliographical References

Chow, S.C.; Liu, J-P. Design and Analysis of Bioavailability and Bioequivalence Studies. New York: Marcel Dekker. 2000

Diletti, E.; Hauschke, D.;Steinijans, V.W. Sample Size Determination for Bioequivalence Assessment By Means of Confidence Intervals, Int. J. Clin. Pharmacol. Therap., 29:1-8. 1991

Guidence for industry - Statistical Approaches to Establishing Bioequivalence
U.S. Department of Health and Human Services; FDA - CDER, January 2001.

Hauck, W.W.; Anderson, S. Types of Bioequivalence and Related Statistical Considerations. Int. J. Clin. Pharmacol. Therap., 30:181-7, 1992.

Liu, J-P. Use of the Repeated Crossover Designs in Assessing Bioequivalence, Stat. Med., 14:1067-78, 1995.

Schuirmann, D.J. Treatment of Bioequivalence Data: Log Transformation, in Proceedings of Bio-International' 89 - Issues in the Evaluation of Bioavailability Data, Toronto, Canada, October 1-4, 159-61, 1989.

Westlake, W.J. The Design and Analysis of Comparative Blood-Level Trials, in Current Concepts in the Pharmaceutical Sciences, Dosage Form Design and Bioavailability (J.Swarbrick, ed.), Lea and Febiger, 149-79, 1973.

Westlake, W.J. Bioavailability and Bioequivalence of Pharmaceutical Formulations, in Biopharmaceutical Statistics for Drug Development (K.E.Peace, ed.), Marcel Dekker, Inc., 329-52, 1988.