BIOAVAILABILITY AND BIOEQUIVALENCE

By H. S Steyn,
Statistical Consultation Services, North-West University (Potchefstroom Campus)

1. **Bioavailability** (see Westlake, 1988)

1.1 Absorption:

The aim is to get the active ingredient of a drug to the site of action in the human body. Since this site (often an organ or tissue) is usually not amenable to sampling, the degree of absorption in the bloodstream can be measured by taking blood samples and determining the concentration of active ingredient over time. The drug can be administered in a number of ways, mainly orally (extravascular) or by injection (intravenous or intramuscular).

1.2 Elimination:

This is the removal of the active ingredient from the body. The major routes of elimination are from the bloodstream via the kidney into urine and fecal excretion.

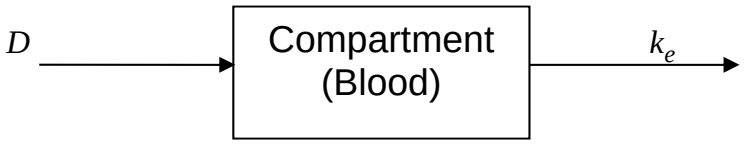
1.3 Blood-level trials:

One dose of a drug formulation is administered to each of *n* subjects (patients or volunteers) and then a series of blood samples are taken from each subject. These samples are then assayed for active ingredient content. The results for each subject is a sequence of blood drug concentrations over time. A typical blood-level profile for one subject can be displayed as in Figure 1. The choice of sampling times is critical if an accurate characterization of the time course of the concentrations is required. More frequent samples are required in the steep ascent up to the peak blood level and at the region of peak levels, while longer time intervals will suffice in the decaying portion of the curve. Sampling also ought to be continued until the concentration becomes negligible, usually more than 3-5 times the estimated half-life of the ingredient content, but, according to EMEA Guidline (Committeefor Medicinal Products for Human Use, 20 January 2010), not longer than 72 hours.

1.4 Pharmacokinetic models:

These are mathematical representation of human organs as a series of compartments with linear transfers between them. We discuss the simplest and most frequently encountered examples of such models:

(a) One compartment model, intravenous:

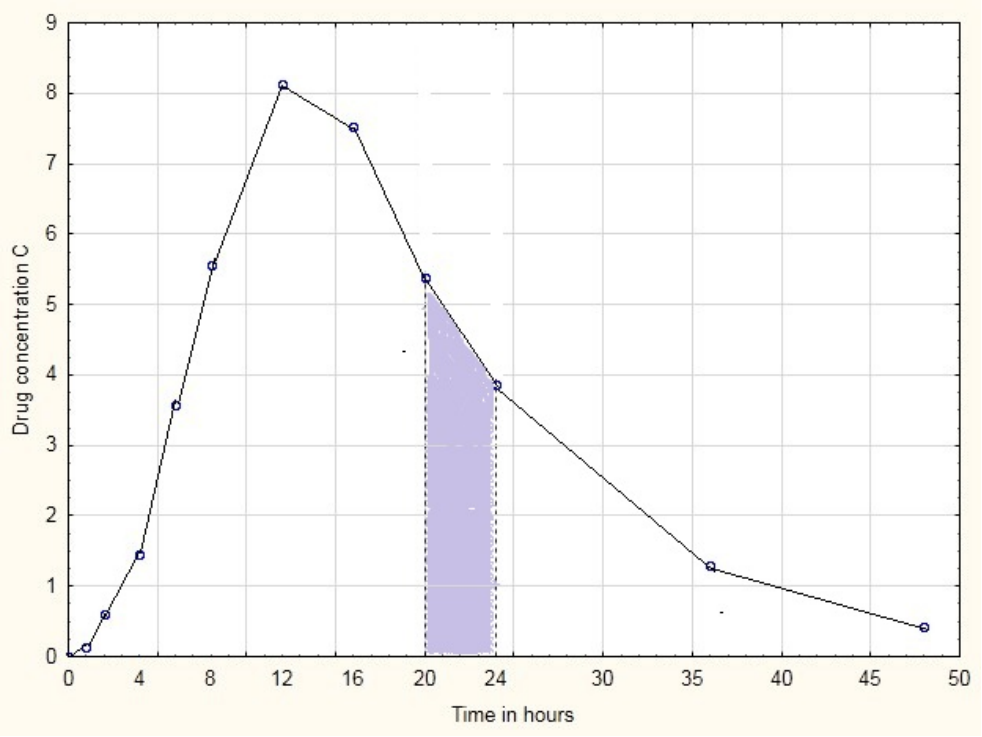


Here dose D is injected instantaneously into the bloodstream at time zero, and k_e is the elimination rate of the active ingredient. By means of a differential equation over time, the blood concentration as function over time can by modelled as:

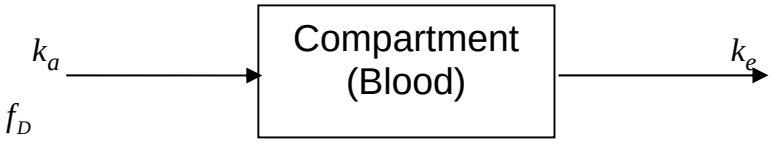
$$Y\left(t_i\right) = C_0 e^{-k_e t_i}\,, \tag{1}$$

where $t_i, i = 1, 2, \ldots, \mathrm{I}$ are the time points and C_0 is the initial concentration injected at time zero.

Figure 1: Blood-level profile of one subject



(b) One compartment, extravascular (or open) model:

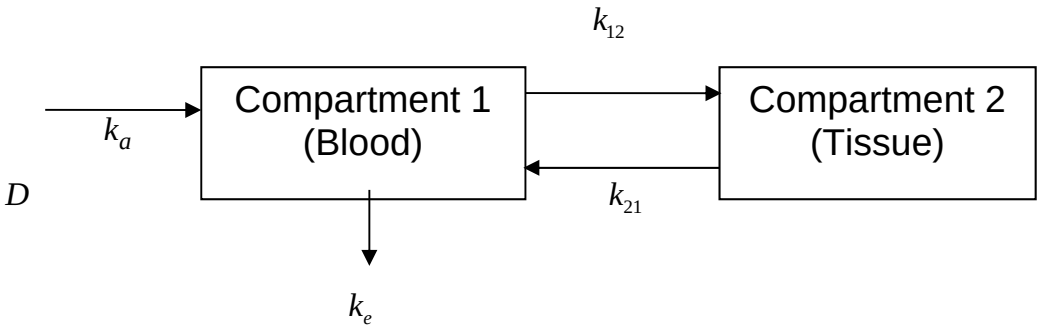


Here an amount A of active ingredient is administered orally and then absorbed into the bloodstream, from which elimination occurs as before. If V is the volume of distribution and f_D is the amount of drug which is absorbed into the bloodstream, then $A = f_D / V$. Let k_a be the absorption rate constant. The differential equation in this case results in:

$$Y(t_i) = A \cdot \frac{k_a}{k_a - k_e} \cdot \left(e^{-k_e t_i} - e^{-k_a t_i} \right), \qquad (2)$$

where it is assumed that $k_a > k_e$.

(c) Two compartment, extravascular (or open) model:



Here the rate constants k_{12} and k_{21} give the rates from the blood to tissue and back. This model will not be discussed this any further in this text.

1.5 Definition of Bioavailability:

It is the extent and rate of which a substance or its therapeutic moiety is delivered from a pharmaceutical form (i.e. a tablet, capsule, injection, etc.) into the general circulation (Wijnand, 1992).

In the case of compartment models, where the blood concentration curve over time is given by $Y(t)$, the bioavailability are given by the following characteristics:

(a) To measure extent: the area under the curve (AUC):

- $AUC(0-T) = \int_0^T Y(t)dt$,

- $AUC(0-\infty) = \int_0^\infty Y(t)dt$,

where T is the last sample's time point (in the example above T=48).

(b) To measure rate:

- Half-life $(t_{½})$: the time required for the blood concentration to fall to half its value at time zero (when the drug was administered intravenously). For one-compartment models, this can be calculated as:

$$t_{½} = \ln 2 / k_e . \qquad (3)$$

- Peak time: The time, $t_{\max}$, associated with the maximum concentration, $C_{\max}$. It can be calculated by solving the following equation in t:

$$\frac{dY(t)}{dt} = 0 . \qquad (4)$$

1.6 Absolute bioavailability:

It is the AUC for extravascular administration (with dosage D_{ev} relative to the AUC of intravenous administration, with dosage D_{iv}): (Wijnand, 1992):

$$F = \frac{AUC_{ev}\left(0-\infty\right)}{AUC_{iv}\left(0-\infty\right)} \cdot \frac{D_{iv}}{D_{ev}} . \qquad (5)$$

1.7 Compartmental bioavailability:

Given a series of measured concentrations over time, the above pharmacokinetic models can be fitted to data to obtain estimates of the pharmacokinetic constants k_e, C_0, k_a, A, k_{12} and k_{21}. Nonlinear regression methods can be used to fit these models, but the estimates of the abovementioned constants and the estimated concentrations $Y\left(t\right)$, may be very different from the true values. This is usually due to large variability and few data points. However, according to the European Medicines Agency (EMEA) 'Guideline on the Investigation of Bioequivalence' (see Committee for Medicinal Products for Human Use, 20 January 2010), this method is not acceptable for the estimation of pharmacokinetic characteristics.

1.8 Non-compartmental bioavailability:

In practice it is simpler to estimate the bioavailability for each subject by directly using its blood-level profile (as in the example).

(a) AUC is calculated by means of the trapezoidal rule (Wijnand, 1992); where $C_i, i = 1,\ldots,\mathrm{I}$ are the measured concentrations at t_i:

- $$AUC\left(0-T\right)_{lin} = ½\sum_{i=2}^{\mathrm{I}}\left[\left(t_i - t_{i-1}\right)\cdot\left(C_i + C_{i-1}\right)\right]. \qquad (6)$$

For example, the area of the trapezium indicated on the blood-concentration profile of Figure 1 is $½\left(t_{10}-t_9\right)\cdot\left(C_{10}+C_9\right) = ½\left(24-20\right)\cdot\left(3.86+5.39\right) = 18.5$. Equation (6) gives the accumulation of areas of all such trapezia in the figure. Here it is assumed that the $Y\left(t\right)$-curve is linear between time-points (as displayed in Figure 1).

However, when one makes the assumption that the $\ell\mathrm{n}\,Y\left(t\right)$-curve is linear between time-points (which is approximately the case for larger t_i - values), the calculation becomes (Wijnand, 1992):

- $$AUC\left(0-T\right)_{\log} = ½\sum_{i=2}^{\mathrm{I}}\left[\left(t_i - t_{i-1}\right)\left(C_i - C_{i-1}\right)/\,\ell\mathrm{n}\left(\frac{C_i}{C_{i-1}}\right)\right]. \qquad (7)$$

The area between T and ∞ can be taken as

$$\int_{\mathrm{T}}^{\infty} Y(t)\,dt = \left[\frac{1}{k_e} Y(t)\right]_{t=T}^{t=\infty} = Y(T)/k_e = C_T \Big/ k_e ,$$

and therefore,

- $$AUC(0-\infty) = AUC(0-T) + C_T \Big/ k_e . \qquad (8)$$

To obtain k_e from the blood-level profile, note that for the <u>one compartment intravenous</u> model:

$$\ln Y(t) = K_1 - k_e t,$$

where K_1 is some constant and therefore $-k_e$ is the slope of the least squares fit of the linear relationship between $\ln(C_i)$ and t_i.

For the <u>one compartment extravascular</u> model, note that for larger t (usually $t \geq t_{max}$),

$$Y(t) = K_2 e^{-k_e t} ,$$

where K_2 is some constant, and $-k_e$ is again the slope of the linear relationship between $\ln(C_i)$ and larger values of t_i.

According to the EMEA guidelines, k_e is a reliable estimate when $AUC(0-T) \geq 0.8\,AUC(0-\infty)$ and there are at least 3-4 time points in the elimination phase.

- $C_{\max}$ is the <u>maximum measured concentration</u> obtained from the blood-level values.
- $t_{\max}$ is the time after administration of the drug that corresponds to the time-point of $C_{\max}$.
- $t_{½}$ is calculated from (3) using the estimated value of k_e.

<u>Example 1:</u>

Serum theophylline concentrations following a single dose of 600 mg were determined on a subject at 20:00 on day 1 and then again at 21:00, 22:00, 24:00. On day 2 the concentrations were determined at 02:00, 04:00, 08:00, 12:00, 16:00, and 20:00, and on day 3 at 08:00 and 20:00. The time (in hours) after administration (t) and the concentrations $C\,(mg/\ell)$ are displayed in Table 1.

The lower part of the table gives the pharmacokinetic characteristics for this subject's blood-level profile. $C_{\max}$ is the maximum of the *C*-values in the second column, while $t_{\max}$ is the associated *t* -value. To calculate AUC, the third column gives the areas for consecutive time intervals (assuming $Y(t)$ is linear between these points) and the sum therefore results in the $AUC(0-48)_{lin}$ value (following formula (6)). The fourth column gives the areas per interval using formula (7), with its total $AUC(0-48)_{\log}$. The

negative of the slope of the linear regression line of $ln\,C$ values (fifth column) with their associated times *t* gives k_e. This value is used to calculate $t_{1/2}$ and $AUC(0-\infty)_{lin}$ and $AUC(0-\infty)_{\log}$. Note that the slope was only determined by using data points in the downward part of the profile (usually the points from $(t_{\max}, C_{\max})$ onwards).

Table 1: Data of times with concentrations of one subject, pharmacokinetic characteristics derived from data.

Hours after dose (t)	Serum concen-tration (C)	Area per interval (linear)	Area per interval (log)	ln C
0	0*			
1	0.14	0.07	0.07#	-1.966
2	0.6	0.37	0.32	-0.511
4	1.45	2.05	1.93	0.372
6	3.57	5.02	4.71	1.273
8	5.57	9.14	8.99	1.717
12	8.14	27.42	27.10	2.097
16	7.53	31.34	31.32	2.019
20	5.39	25.84	25.60	1.685
24	3.86	18.5	18.33	1.351
36	1.29	30.9	28.14	0.255
48	0.42	10.26	9.30	-0.868

* value below detection limit, taken as 0

use linear area, log not defined when C = 0

Cmax	8.14
tmax	12
AUC(0-48)lin	160.9
AUC(0-48) log	155.8
k_e	0.086
$t_{½}$	8.08
AUC(0-inf)lin	165.8
AUC(0-inf) log	160.7

Note: The EXCEL spreadsheet which was employed to perform most of the calculations shown here is attached to this document as supplementary material.

2. Bioequivalence

Birkett (2003) defined bioequivalence by stating that, "*two pharmaceutical products are bioequivalent if they are pharmaceutically equivalent and their bioavailabilities (rate and extent of availability) after administration in the same molar dose are similar to such a degree that their effects, with respect to both efficacy and safety, can be expected to be essentially the same. Pharmaceutical equivalence implies the same amount of the same active substance(s), in the same dosage form, for the same route of administration and meeting the same or comparable standards.*"

2.1 Hypothesis of Bioequivalence (see Steinijans & Hanschke, 1990):

Two drug formulations whose rate and extent of absorption differ by a prescribed 100 *K*% or less, are generally considered as equivalent. The Food and Drug Administration (FDA) requires this percentage to be at most 20%. Let μ_T and μ_R be the expected means or medians of characteristics like the AUC, or $C_{\max}$, etc., of a test (T) and a reference (R) formulation, respectively. We then consider *T* and *R* to be bioequivalent if

$$|\mu_T - \mu_R| < K\mu_R$$

or

$$1-K < \mu_T/\mu_R < 1+K\,.$$

Generally bioequivalence is then obtained if

$K_1 < \mu_T/\mu_R < K_2$, where

$0 < K_1 < 1 < K_2\,.$

Usually $K_1 = 0.8$, $K_2 = 1.25$, since 1.25/1 = 1/0.8 (see FDA Guidance, Centre for Drug Evaluation and Research, 2003, as well as EMEA Guideline, Committee for Medicinal Products for Human Use, 20 January 2010). The values of $K_1 = 0.7$ and $K_2 = 1.43$ have also been suggested, especially for $C_{\max}$ in fast releasing drugs.

Biostatisticians have repeatedly pointed out that the conventional null-hypothesis $\mu_T = \mu_R$ is not appropriate in bioequivalence testing. Since the primary concern here is the protection of the patient (i.e. the consumer) against the acceptance of bioequivalence if it is not true. Therefore, to limit this consumer risk of erroneously accepting bioequivalence, the alternative hypothesis (H_1) should be bioequivalence and bioinequivalence has to be formulated as the null-hypothesis (H_0).

Therefore:

$$H_0 : \mu_T / \mu_R \le K_1 \quad or \quad \mu_T / \mu_R \ge K_2$$
$$H_1 : K_1 < \mu_T / \mu_R < K_2 \qquad (9)$$

2.2 Testing procedures of bioequivalence:

Split up the hypotheses in (9) for two one-sided tests:

$$H_{01} : \mu_T / \mu_R \le K_1 \quad \text{against} \quad H_{11} : \mu_T / \mu_R > K_1$$

and

$$H_{02} : \mu_T / \mu_R \ge K_2 \quad \text{against} \quad H_{12} : \mu_T / \mu_R < K_2$$

The rejection of both H_{01} and H_{02} at a $100\alpha\%$ level is equivalent to the inclusion of the shortest $100(1-2\alpha)\%$ confidence interval (CI) for μ_T / μ_R within the bioequivalence range (K_1, K_2).

The procedure in the case of the bioequivalence range (0.8,1.2) and $\alpha = 0.05$ (i.e. using a 90% confidence interval for θ= μ_T / μ_R) is displayed in Table 2 (see Steyn et al.,1988).

In the second part of the table a method is proposed where one can conclude that no decision concerning bioequivalence can be made; this opens the possibility of repeating the bioequivalence trial (e.g. by using more subjects).

2.3 The two-way crossover design (see Steinijans & Diletti, 1983):

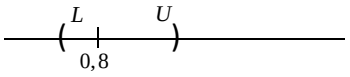
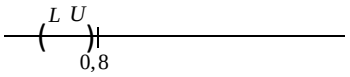
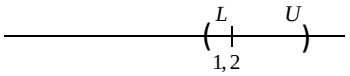
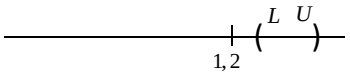
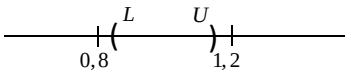
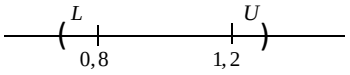
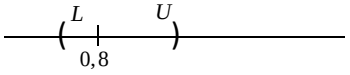
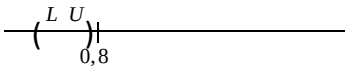
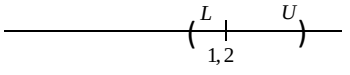
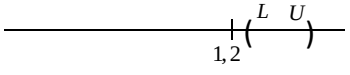
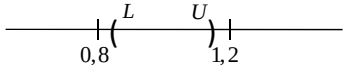
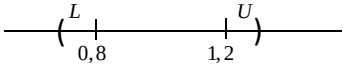
This design enables the researcher to estimate the treatment (i.e. *T* vs. *R*) effect by controlling for subject and phase effect. A random sample of $n/2$ subjects is drawn from the n (even number) available subjects (sequence 1-group) and the test formulation is administered to them. According to the EMEA guideline (Committee for Medicinal Products for Human Use, 20 January 2010), the requirement is $n \ge 12$. After a washout-period they receive the reference formulation. For the remaining $n/2$ subjects (sequence 2 group) the sequence of administration of the formulations is switched around.

Table 3 displays the expected values y_{ijk} of the measurements, $i = 1,2$ (sequences), $j = 1,2$ (phases) and $k = 1,2,\ldots,n/2$ (subjects). In the table μ is the overall mean, π_j the j^{th} phase effect,

and τ_ℓ is the ℓ^{th}, (ℓ = 1, 2) treatment effect (all fixed effects). A random effect s_{ik} is associated for each subject $k = 1, 2, \ldots, n/2$ within the i^{th} sequence.

Note also that it is assumed that the carryover effect is zero. This is accomplished by ensuring that a sufficiently long washout-period of 5-6 half-lives is used to make certain that no measurable residual drug is carried over from the first to the second phase.

Table 2: Schematic presentation of the possibilities of rejecting or accepting bioequivalence using the interval and the adapted interval methods.

Hypothesis	Test	Possibilities in terms of the confidence limits (L,U)*	Criteria	Decision
			$L \le 0,8$	Reject bioequivalence
				
Interval method			$U \ge 1,2$	Reject bioequivalence
				
$H_1 : 0,8 < \theta < 1,2$			$L > 0,8; \mathrm{U} < 1,2$	Accept bioequivalence
$H_0 : \theta \le 0,8$			$L < 0,8; U > 1,2$	Reject bioequivalence
or			$L \le 0,8 \le U$	No decision
$\theta \ge 1,2$			$U < 0,8$	Reject bioequivalence
Proposed method			$L \le 1,2 \le U$	No decision
			$L > 1,2$	Reject bioequivalence
			$L > 0,8; U < 1,2$	Accept bioequivalence
			$L < 0,8; U > 1,2$	No decision

* L,U is respectively the lower and upper limits of the 90% confidence interval for θ.

Denote the measurements as:

$$Y_{ijk} = y_{ijk} + e_{ijk},$$

where e_{ijk} are the error terms which are $N\left(0, \sigma_e^2\right)$ distributed. Further, denote $\overline{Y}_{ij\bullet}$ as the mean over the measurements in the i^{th} sequence and j^{th} phase.

Table 3: Expected values y_{ijk} in a two-way crossover design, $k = 1, 2, \ldots, n/2$.

	Phase 1	Phase 2
Sequence group 1	$y_{11k} =$ $\mu + \pi_1 + \tau_1 + s_{1k}$	$y_{12k} =$ $\mu + \pi_2 + \tau_2 + s_{1k}$
Sequence group 2	$y_{21k} =$ $\mu + \pi_1 + \tau_2 + s_{2k}$	$y_{22k} =$ $\mu + \pi_2 + \tau_1 + s_{2k}$

Note that other designs can also be used: (a) a parallel design (i.e. only one phase) when $t_{½}$ is large implying a very long washout period, and (b) a crossover design with replicates (for example two phases for each product), which have the following advantages (see Centre for Drug Evaluation and Research, 2003): (1) allow comparisons of within-subject variances for the test and reference products, (2) provides more information about the intrinsic factors underlying formulation performance, and (3) reduces the number of subjects participating in the BE study.

2.4 Test for carry-over effect:

To estimate the sequence effect, the means over all subjects (for both phases) in sequence 2 can be subtracted from the means over all subjects for both phases in sequence 1:

$$\frac{1}{2}\left(\overline{Y}_{11\bullet} + \overline{Y}_{12\bullet} - \overline{Y}_{21\bullet} - \overline{Y}_{22\bullet}\right).$$

Testing for a carry-over effect is equivalent to testing for a sequence effect (see Cotton, 1989). This can be done by an analysis of variance (ANOVA) with sequence as the main effect and subjects within sequences as the error effects (i.e. the interaction between the random subject and the sequence effects). If this analysis results in a non-significant sequence effect, no carry-over effect can be concluded.

Note that the EMEA guideline (Committee for Medicinal Products for Human Use, 20 January 2010) states: "*If there are any subjects for whom the pre-dose concentration is greater than 5 percent of the* $C_{\max}$ *value for the subject in that period, the statistical analysis should be performed with the data from that subject for that period excluded. In a 2-period trial this will result in the subject being removed from the analysis. The trial will no longer be considered acceptable if these exclusions result in fewer than12 subjects being evaluable*".

2.5 Estimation of the treatment effect:

Take $\bar{D}_{\bullet\bullet} = \frac{1}{2}\left(\bar{Y}_{11\bullet} + \bar{Y}_{22\bullet}\right) - \frac{1}{2}\left(\bar{Y}_{12\bullet} + \bar{Y}_{21\bullet}\right)$, (10)

then from Table 3 it follows that $E\left(\bar{D}_{\bullet\bullet}\right) = \tau_1 - \tau_2$, the difference in treatment effects. Under the assumption of normality, the $100(1-2\alpha)\%$ confidence interval for $\tau_1 - \tau_2$ is:

$$\bar{D}_{\bullet\bullet} \pm t\left(n-2, 1-2\alpha\right)\sqrt{2MSE/n}\,, \tag{11}$$

where *MSE* is the mean square error in the analysis of variance (ANOVA) with fixed effects for treatment and phase and random effect for subjects, but without any interactions. The term $t(n-2,1-2\alpha)$ represents the $(1-2\alpha)^{th}$ quantile of the t - distribution with $n-2$ degrees of freedom.

Note that if the sequence group sizes are n_1 and n_2 instead of $n/2$, then (11) becomes

$$\bar{D}_{\bullet\bullet} \pm t\left(n_1+n_2-2, 1-2\alpha\right)\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)MSE/2}\,. \tag{12}$$

However, if a significant sequence effect (i.e. carryover effect) exists, only the first phase's data can be used. The treatment effect can then be estimated by

$$\bar{D}_{\bullet\bullet} = \bar{Y}_{11\bullet} - \bar{Y}_{21\bullet}\,, \tag{13}$$

with

$$\frac{\left(n_1-1\right)S_1^2 + \left(n_2-1\right)S_2^2}{n_1+n_2-2}\,, \tag{14}$$

substituted for $MSE/2$ in equation (12) and where n_i and S_i are the sizes and standard deviations of the two sequence groups in phase 1, respectively.

2.6 Ratios and log-transformation:

Since we rather need a $100(1-2\alpha)\%$ CI for μ_T/μ_R and not for $\mu_T-\mu_R$ as equation (11) suggests, take $\tau_T=\ell n\mu_T$ and $\tau_R=\ell n\mu_R$, so that

$$\tau_T-\tau_R=\ell n(\mu_T/\mu_R)$$

or

$$\mu_T/\mu_R=e^{\mu_T-\mu_R}\ . \qquad (15)$$

Taking the confidence interval in (11) as (L,U), the $100(1-2\alpha)\%$ CI for μ_T/μ_R is given by (e^L,e^U).

Let X_{ijk} represent measurements like the AUC or the value $C_{\max}$ etc., and define $Y_{ijk}=\ell nX_{ijk}$, then

$$e^{\bar{D}_{\square\square}}=\exp\left[\left\{½\sum_{k=1}^{n/2}\left(\ell nX_{11k}-\ell nX_{12k}\right)/\left(\tfrac{n}{2}\right)\right\}\left\{½\sum_{k=1}^{n/2}\left(\ell nX_{22k}-\ell nX_{21k}\right)/\left(\tfrac{n}{2}\right)\right\}\right]$$

$$=\sqrt{\prod_{k=1}^{n/2}\left(X_{11k}/X_{12k}\right)^{2/n}}\times\sqrt{\prod_{k=1}^{n/2}\left(X_{22k}/X_{21k}\right)^{2/n}}$$

$$=\sqrt{GM_1\times GM_2}\ , \qquad (16)$$

where GM_1 is the geometric mean of the within subject ratio of *T* relative of *R* for the first sequence and GM_2 that for the second sequence. Therefore $\sqrt{GM_1\times GM_2}$, which again is the geometric mean of the two sequence's geometric means, can be used as an estimator for μ_T/μ_R, since

$$E\left(e^{\bar{D}_{\bullet\bullet}}\right)=e^{\tau_T-\tau_R}\ .$$

Note that (16) also holds if the sequence group sizes are n_1 and n_2 instead of $n/2$.

2.7 A nonparametric confidence interval for $\mu_T-\mu_R$:

In the parametric CI above, the assumption of normality for the logarithms of the measurements like AUC, $C_{\max}$, and $t_{½}$ has to be made. Hauschke et al. (1990) suggested the following nonparametric method when normality does not necessarily hold:

Let $Y_{ij}=Y_{i1j}-Y_{i2j}$, be the intra individual differences of the logarithmic transformed measurements for the first and second phases for sequence group *i*, where $j=1,2,\ldots,n_i$.

The two-phase crossover design can be reduced to a two-sample situation concerning the treatment sequences T/R and R/T, which differ only by the shift parameter $2\ell n(\pi_1/\pi_2)=2\ell n(\mu_T/\mu_R)$.

A two-sample test can be performed with the following two sets of one-sided hypotheses:

$$H_{01}: \ell n(\mu_T / \mu_R) \le \ell n K_1 \; against \; H_{11}: \ell n(\mu_T / \mu_R) > \ell n K_1$$

and

$$H_{02}: \ell n(\mu_T / \mu_R) \ge \ell n K_2 \; against \; H_{12}: \ell n(\mu_T / \mu_R) < \ell n K_2.$$

As before, the rejection of H_{01} and H_{02} by nonparametric Mann-Whitney tests each at level α, is equivalent to the inclusion to the corresponding nonparametric $100(1-2\alpha)\%$ CI for $\mu_T - \mu_R$ in the bioequivalence range (K_1, K_2).

The procedure is the following:

a) From the n_1 differences Y_{1j} between the first and second phases in the treatment sequence *T/R* and the n_2 differences in the treatment sequence *R/T*, the $n_1 n_2$ pairwise difference $Y_{1j} - Y_{2j*}$ $(j = 1,\ldots,n_1,\ j* = 1,\ldots,n_2)$ are calculated.
b) Rank these differences according to magnitude.
c) The median to these differences serves as Hodges-Lehmann point estimator for $2\ell n(\mu_T / \mu_R)$ and if it is divided by 2 and the exponent is taken, this will estimate $\mu_T - \mu_R$.
d) Table 2 from Hauschke et al. (1990) gives the indices corresponding to the ranked values $Y_{1j} - Y_{2j*}$ for given n_1 and n_2 values ≤ 12. These differences divided by 2 give the lower and upper 90% (in the case of $\alpha = 0.05$) confidence limits for $\ell n(\mu_T / \mu_R)$ (The table is reproduced here as Table 4).
e) Exponentially transformed lower and upper limits give the 90% confidence limits for $\mu_T - \mu_R$.
f) For $n_1, n_2 > 12$, but $n_1 \le 40$ and $n_2 \le 20$ other α values the indices are determined as follows:
   - Use the tables of Milton (Milton, 1964) (see the PDF-document 'Tables of Mann-Whitney critical values' as appendix to this document).
   - Read the value $k = u$ from the tables at $n_1 = m$ and $n_2 = n$ and $P_r(U_{y>x} \le u) \le \alpha$.
   - The index for the lower limit of the $100(1-\alpha)\%$ CI for $\ell n(\mu_T / \mu_R)$ is k+1 (e.g. for $n_1 = 12, n_2 = 12, \alpha = 0.05, k = 42$, so that the index is 43).
   - The index for the upper limit is $n_1 n_2 - k$ (in the example above 12x12-42=102).

g) When either $n_1 > 40$ or $n_2 > 20$, the asymptotic normality of the Mann-Whitney U-statistic can be utilised as follows:

$$k = \left[-z_\alpha \sqrt{n_1 n_2 \left(n_1 + n_2 + 1\right)/12} + \frac{1}{2} n_1 n_2 \right], \qquad (17)$$

where z_α is the standard normal $\left(1-\alpha\right)100^{th}$ quantile (e.g. for $\alpha = 0.05$, $z_\alpha = 1.645$), and $\left[x\right]$ is the largest integer $\leq x$.

For example when $n_1 = 25$ and $n_2 = 30$, then

$$k = \left[-1.645\sqrt{25 \times 30 \times 56/12} + \frac{1}{2}(25 \times 30) \right]$$

$$= \left[277.68\right] = 277 .$$

2.8 Bioequivalence for $t_{\max}$

The EMEA guideline (Committee for Medicinal Products for Human Use, 20 January 2010) states: "*A statistical evaluation of $t_{\max}$ is not required. However, if rapid release is claimed to be clinically relevant and of importance for onset of action or is related to adverse events, there should be no apparent difference in median $t_{\max}$ and its variability between test and reference product.*"

Table 4: Indices for the construction of 90%-confidence intervals. n_1 and n_2 denote the numbers of subjects in the respective treatment sequences test/reference and reference/test. Listed below these indices is the exact confidence coefficient.

n_1	n_2=4	5	6	7	8	9	10	11	12
4	2/15	3/18	4/21	5/24	6/27	7/30	8/33	9/36	10/39
	0.9429	0.9365	0.9333	0.9273	0.9273	0.9245	0.9241	0.9223	0.9220
5	3/18	5/21	6/25	7/29	9/32	10/36	12/39	13/43	14/47
	0.9365	0.9048	0.9177	0.9268	0.9068	0.9171	0.9008	0.9103	0.9182
6	4/21	6/25	8/29	9/34	11/38	13/42	15/46	17/50	18/55
	0.9333	0.9177	0.9069	0.9266	0.9187	0.9121	0.9066	0.9017	0.9169
7	5/24	7/29	9/34	12/38	14/43	16/48	18/53	20/58	22/63
	0.9273	0.9268	0.9266	0.9027	0.9061	0.9093	0.9122	0.9147	0.9169
8	6/27	9/32	11/38	14/43	16/49	19/54	21/60	24/65	27/70
	0.9273	0.9068	0.9187	0.9061	0.9170	0.9073	0.9169	0.9092	0.9021
9	7/30	10/36	13/42	16/48	19/54	22/60	25/66	28/72	31/78
	0.9245	0.9171	0.9121	0.9093	0.9073	0.9061	0.9053	0.9048	09045
10	8/33	12/39	15/46	18/53	21/60	25/66	28/73	32/79	35/86
	0.9241	0.9008	0.9066	0.9122	0.9169	0.9053	0.9108	0.9014	0.9069
11	9/36	13/43	17/50	20/58	24/65	28/72	32/79	35/87	39/94
	0.9223	0.9103	0.9017	0.9147	0.9092	0.9048	0.9014	0.9121	0.9092
12	10/39	14/47	18/55	22/63	27/70	31/78	35/86	39/94	43/102
	0.9220	0.9182	0.9169	0.9169	0.9021	0.9045	0.9069	0.9092	0.9113

Example 2 (Steinijans & Diletti, 1983):

Twelve healthy volunteers participated in a crossover study to investigate the influence of food intake on bioavailability of theophylline from a sustained-release aminophylline preparation. The case in which the drug was taken after fasting overnight and a standard breakfast was eaten 2 hours after taking the drug, serves as the reference (*R*). Drug intake directly after consumption of the same standard breakfast is the test situation (*T*). Serum theophylline levels were determined before and at 0.5, 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 24 and 32 hours after administration of drugs.

Table 5: AUC-values per subject for Reference and Test formulations.

Subject	Sequence	Reference (R)	Test (T)	Ratio T/R	ln(T/R)
1	T/R	136.0	135.7	0.998	-0.0010
2	T/R	152.6	155.3	1.018	0.0076
3	R/T	123.1	148.9	1.210	0.0826
4	R/T	77.0	81.2	1.055	0.0231
5	T/R	115.7	139.2	1.203	0.0803
6	T/R	72.0	91.7	1.274	0.1050
7	R/T	116.4	118.7	1.020	0.0085
8	T/R	151.1	133.2	0.882	-0.0548
9	R/T	118.9	115.6	0.972	-0.0122
10	T/R	156.1	150.3	0.963	-0.0164
11	R/T	222.4	223.9	1.007	0.0029
12	R/T	158.1	154.1	0.975	-0.0111

Table 5 displays the AUC-values for *R* and *T* for each subject as well as the ratio R/T and $\ell n\left(T/R\right)$. Using the sequences in which subjects received the treatments, Table 6 was constructed as data for the analyses of variance (ANOVAs) in Tables 7 and 8.

Table 7 displays the results of an ANOVA where the sequence effect is tested for significance. This was done by using the error mean square for the random subject effect which was nested within sequence (i.e. the sequence by subject interaction, *sequence* subject*). Since p=0.93, no sequence effect existed, from which the conclusion can be drawn that no carry-over between phases occurred.

Finally, a 3-way ANOVA with main effects treatment, phase and subject (random) and $\ell n\left(AUC\right)$ as dependent variable, was performed. The ANOVA results are displayed in Table 8, while the estimated treatment difference $\bar{D}_{\bullet\bullet}$ and its standard error are given in Table 9 as 0.018 and 0.014.

Table 6: Data of AUC and ln(AUC) values for analyses of variance (ANOVAs).

Subject	Sequence	Treatment	Phase	AUC	ln(AUC)
1	T/R	R	2	136	2.134
2	T/R	R	2	152.6	2.184
3	R/T	R	1	123.1	2.090
4	R/T	R	1	77	1.886
5	T/R	R	2	115.7	2.063
6	T/R	R	2	72	1.857
7	R/T	R	1	116.4	2.066
8	T/R	R	2	151.1	2.179
9	R/T	R	1	118.9	2.075
10	T/R	R	2	156.1	2.193
11	R/T	R	1	222.4	2.347
12	R/T	R	1	158.1	2.199
1	T/R	T	1	135.7	2.133
2	T/R	T	1	155.3	2.191
3	R/T	T	2	148.9	2.173
4	R/T	T	2	81.2	1.910
5	T/R	T	1	139.2	2.144
6	T/R	T	1	91.7	1.962
7	R/T	T	2	118.7	2.074
8	T/R	T	1	133.2	2.125
9	R/T	T	2	115.6	2.063
10	T/R	T	1	150.3	2.177
11	R/T	T	2	223.9	2.350
12	R/T	T	2	154.1	2.188

For $\alpha = 0.05$, a 90% CI is required for the expected treatment difference $\tau_1 - \tau_2$.

Since $t(10, 0.9) = 1.812$, the confidence limits are

$$0.018 \pm 1.812 \text{ x } 0.014 = (-0.007, 0.043),$$

and according to Table 9: (-0.008, 0.044).

In terms of a ratio, τ_T / τ_R is estimated by $e^{\bar{D}_{\bullet\bullet}} = e^{0.018} = 1.018$, while the 90% CI is (0.992,1.045). Clearly this 90% CI is contained within (0.8, 1.25) and therefore bioequivalence can be concluded.

Table 7: The SPSS output of an ANOVA to test for a carry-over effect.

Dependent Variable: ln(AUC)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	107.371	1	107.371	3231.912	.000
	Error	.332	10	.033[a]		
Sequence	Hypothesis	.000	1	.000	.008	.930
	Error	.332	10	.033[a]		
Sequence * Subject	Hypothesis	.332	10	.033	27.952	.000
	Error	.014	12	.001[b]		

a. MS(Sequence * Subject)

b. MS(Error)

Table 8: The SPSS output of a 3-way ANOVA with main effects Treatment, Phase and Subject.

Dependent Variable: ln(AUC)

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	107.371	1	107.371	3552.250	.000
	Error	.332	11	.030[a]		
Treatment	Hypothesis	.002	1	.002	1.558	.240
	Error	.012	10	.001[b]		
Phase	Hypothesis	3.044E-05	1	3.044E-05	.025	.878
	Error	.012	10	.001[b]		
Subject	Hypothesis	.332	11	.030	24.547	.000
	Error	.012	10	.001[b]		

a. MS(Subject)

b. MS(Error)

Table 9: SPSS output for estimated treatment effect and standard error.

Contrast Results

Treatment Contrast	Difference		Dependent Variable ln(AUC)
Level 2 vs. Level 1	Contrast Estimate		.018
	Hypothesized Value		0
	Difference (Estimate - Hypothesized)		.018
	Std. Error		.014
	Sig.		.240
	90% Confidence Interval for Difference	Lower Bound	-.008
		Upper Bound	.044

The nonparametric estimates can be obtained by using the EXCEL spreadsheet (most of the calculations can be performed in this way) which is attached to this document as supplementary material. The calculations and output from the EXCEL spreadsheet is displayed in Table 10. Here the 6th column gives the intra-individual differences of phases 1 vs 2, while the next 6 columns are the pairwise differences between the 2 sequences' values in column 6. These 36 values are ranked in the last column and indices 1 – 36 are then allocated. From Table 4 the index values 8 and 29 are obtained, from which the values -0.030 and 0.214 are read off from the ranked column. Also, the median is the mean of the 18th and 19th values, i.e. 0.5 (0.051+0.064) = 0.058. The ratio τ_T / τ_R is estimated by $e^{0.058/2} = 1.029$ and the 90% lower limit is given by $e^{-0.030/2} = 0.985$,while the upper limit is $e^{0.214/2} = 1.113,$ and therefore bioequivalence can be concluded.

Note that due to a possible deviation from normality of the ln(AUC) values, the upper CI limit of the ratio differs from those obtained above.

Table 10: Display of the EXCEL spreadsheet used to calculate the nonparametric estimates.

Sub-ject	Se-quen-ce	Refe-rence R	Test T	ln(R)	ln(T)	Dif (Ph1-Ph2)	Y11-Y2j	Y12-Y2j	Y13-Y2j	Y14-Y2j	Y15-Y2j	Y16-Y2j	Index	Ranked(Y1j-Y2j*)
3	R/T	123.1	148.9	4.813	5.003	-0.190	0.188	0.208	0.375	0.432	0.064	0.152	1	-0.154
4	R/T	77.0	81.2	4.344	4.397	-0.053	0.051	0.071	0.238	0.295	-0.073	0.015	2	-0.152
7	R/T	116.4	118.7	4.757	4.777	-0.020	0.017	0.037	0.204	0.261	-0.107	-0.018	3	-0.107
9	R/T	118.9	115.6	4.778	4.750	0.028	-0.030	-0.011	0.157	0.214	-0.154	-0.066	4	-0.073
11	R/T	222.4	223.9	5.404	5.411	-0.007	0.005	0.024	0.192	0.249	-0.119	-0.031	5	-0.066
12	R/T	158.1	154.1	5.063	5.038	0.026	-0.028	-0.008	0.159	0.216	-0.152	-0.063	6	-0.063
1	T/R	136.0	135.7	4.913	4.910	-0.002							7	-0.031
2	T/R	152.6	155.3	5.028	5.045	0.018							8	-0.030
5	T/R	115.7	139.2	4.751	4.936	0.185							9	-0.028
6	T/R	72.0	91.7	4.277	4.519	0.242							10	-0.018
8	T/R	151.1	133.2	5.018	4.892	-0.126							11	-0.011
10	T/R	156.1	150.3	5.050	5.013	-0.038							12	-0.008
													13	0.005
													14	0.015
													15	0.017
													16	0.024
													17	0.037
													18	0.051
													19	0.064
													20	0.071
													21	0.152
													22	0.152
													23	0.157
													24	0.159
													25	0.188
													26	0.192
													27	0.204
													28	0.208
													29	0.214
													30	0.216
													31	0.238
													32	0.249
													33	0.261
													34	0.295
													35	0.375
													36	0.432

References:

Birkett D.J. (2003). Generics - equal or not?. *Aust Prescr* **26**: 85–7.

Centre for Drug Evaluation and Research (2003). "Guidance for Industry: Bioavailability and Bioequivalence Studies for orally Administered Drug Products – General Considerations". (http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegularitoryInformation/Guidances/ucm070124.pdf). United States Food and Drug Administration. Retrieved 17 March 2014.

Committee for Medicinal Products for Human Use (20 January 2010). "Guideline on the Investigation Of Bioequivalence". (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/01/WC500070039.pdf). European Medicines Agency. Retrieved 17 March 2014.

Cotton, John W. (1989). Interpreting data from two-period design (also termed the replicated 2 x 2 Latin square design). *Psychological Bulletin*, 106, 503 – 515.

Hauschke, D., Steinijans, V.W. and Diletti, E. (1990). A distribution-free procedure for statistical analysis of bioequivalence studies. *International Journal of Clinical Pharmacology, Therapy and Toxicology*, 28(2), 72 - 78.

Milton, Roy C. (1964). An Extended Table of Critical Values for the Mann-Whitney (Wilcoxon) Two-Sample Statistic. *Journal of the American Statistical Association*, 59, 925- 934.

Steinijans, V.W. and Diletti, E. (1983). Statistical analysis of bioavailability studies: parametric and nonparametric confidence intervals. *European Journal of Clinical Pharmacology*, 24, 127 – 136.

Steinijans, V.W. and Hauschke, D. (1990). Update on the analyis of bioequivalence studies. *International Journal of Clinical Pharmacology,Therapy and Toxicology*, 28(3),105 - 110.

Steyn, H.S. , Koeleman, H.A. and Bonescans, B. (1988). Testing of bioequivalence of medicines using confidence intervals. South African Journal of Science, 84, 367-371.

Westlake, Wilfred J. (1988). Bioavailability and Bioequivalence of Pharmaceutical Formulations. *Chapter 7 in Biopharmaceutical Statistics for Drug Development*, edited by Karl E. Peace. Marcel Dekker, Inc., New York.

Wijnand, Herman P. (1992). The determination of absolute bioavailability for drug substances with long elimination half-lives (with PC-programs for the method of truncated areas). *Computer Methods and Programs in Biomedicine*, 39, 61-73.