Clinical Trials A Practical Guide to Design, Analysis, and Reporting
Duolao Wang, PhD Ameet Bakhai, MBBS, MRCP

Clinical Trials

A Practical Guide to Design, Analysis, and Reporting

drug x
control

Duolao Wang, PhD
Statistician

Ameet Bakhai, MBBS, MRCP
Cardiologist

REMEDICA

Bioequivalence Trials

Duolao Wang, Radivoj Arezina,
and Ameet Bakhai

Bioequivalence studies are, by and large, conducted to compare a generic drug preparation with a currently marketed formulation – typically an innovator drug product. In a bioequivalence study, blood samples are collected just before the administration of a drug and for a period of time after administration. The drug concentrations are then plotted against time in order to derive pharmacokinetic parameters and evaluate the bioequivalence of the drugs under study. In this chapter, we describe some of the practical issues involved in the design and evaluation of bioequivalence studies and show how inappropriate design of such studies can lead to erroneous conclusions.

What is a bioequivalence trial?

A bioequivalence trial is a study of presumed therapeutic equivalence based on pharmacokinetic (PK) parameters rather than on clinical, or other, endpoints. There are several *in vivo* and *in vitro* methods that can be utilized to evaluate therapeutic equivalence between two medicinal products. In ascending order of preference, these include:

- *in vitro* studies
- comparative clinical studies
- pharmacodynamic (PD) studies
- PK studies

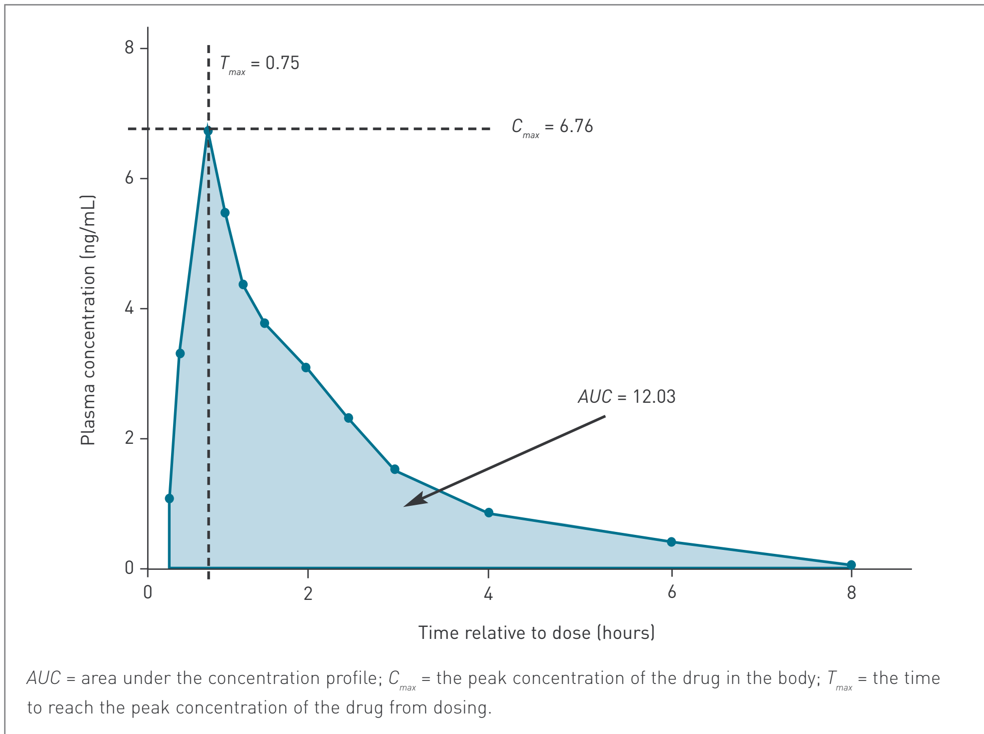
While some of these methods are appropriate only in certain circumstances (eg, *in vitro* dissolution tests can be used to evaluate the therapeutic equivalence of highly soluble, rapidly dissolving, orally active drugs), others (comparative clinical and PD studies) are considered less reliable and are generally only recommended if the PK approach is not possible [1–3]. Quite often, comparisons based on PD endpoints and, in particular, clinical endpoints, prove to be very difficult. Such studies are frequently hindered by factors such as a lack of clearly defined endpoints and huge variability in the measured parameters. Hence, the PK approach is commonly accepted as the method of choice for evaluating therapeutic equivalence between a generic and an innovator (reference) medicinal product. Bioequivalence studies compare PK parameters such as peak concentration (C_{max}) derived from plasma, serum concentrations, and blood concentrations, as described below.

Which PK parameters are used in a bioequivalence study?

To illustrate the PK parameters used, data from an anonymous study of the bioequivalence of generic and reference anagrelide products (used to decrease platelet count) will be used. An anagrelide 'plasma concentration over time' profile of a volunteer is shown in **Figure 1**. The raw PK data are also given in the first three columns of **Table 1**. In a PK study, the following parameters are usually derived from the PK profile in order to describe the drugs studied:

- *AUC*: the area under the 'plasma concentration over time' curve, which describes the total number of drug molecules present in the plasma, thereby providing information on the extent of absorption.

Figure 1. An example of a subject's anagrelide concentration profile.



AUC = area under the concentration profile; C_{max} = the peak concentration of the drug in the body; T_{max} = the time to reach the peak concentration of the drug from dosing.

- C_{max}: the peak concentration of the drug in the body.
- T_{max}: the time, from dosing, to reach C_{max} (C_{max} and T_{max} together are indirect indicators of the rate of absorption).
- λ: the elimination constant, which describes the loss of drug activity from the body per time unit (eg, per hour).
- $T_{1/2}$: the elimination half-life. This is the time required for the amount (or concentration) of the drug in the body to decrease by half.

These parameters fully describe the shape of the 'plasma concentration over time' profile of a study drug. The absorption and elimination phases of a drug are distinguished by the parameters C_{max} and T_{max}. When the amount of the drug absorbed equals the amount eliminated, C_{max} is reached. Before C_{max} is reached, absorption is higher than elimination, and after C_{max} is reached, the situation is reversed.

Table 1. Anagrelide concentration and calculation of *AUC*.

Samples	T_i (h)	C_i (ng/mL)	$(C_i + C_{i-1}) / 2$	$T_i + T_{i-1}$	AUC_i (ng.h/mL)
1	0.00	<LLQ	NA	NA	NA
2	0.25	0.96	NA	0.25	NA
3	0.50	3.19	2.08	0.25	0.52
4	0.75	6.76	4.97	0.25	1.24
5	1.00	5.24	6.00	0.25	1.50
6	1.25	4.20	4.72	0.25	1.18
7	1.50	3.61	3.91	0.25	0.98
8	2.00	2.96	3.29	0.50	1.64
9	2.50	2.24	2.60	0.50	1.30
10	3.00	1.60	1.87	0.50	0.94
11	4.00	0.95	1.28	1.00	1.23
12	6.00	0.29	0.62	2.00	1.18
13	8.00	0.09	0.16	2.00	0.32
14	10.00	<LLQ	NA	2.00	NA
15	12.00	<LLQ	NA	2.00	NA
Total					12.03

AUC = area under the curve; *C* = concentration; LLQ = lower limit of quantitation (0.05 ng/mL); NA = not applicable; *T* = time.

Calculation of PK parameters

C_{max} and T_{max}

The parameters C_{max} and T_{max} can be directly observed from the PK profile for a subject. For the anagrelide data in **Table 1** or **Figure 1**, it is easy to see that C_{max} = 6.76 ng/mL and T_{max} = 0.75 hours.

$AUC_{0\text{–}t}$

$AUC_{0\text{–}t}$ stands for the area under the PK concentration profile from time zero to time t, where t is the last time point at which there is a quantifiable plasma concentration. $AUC_{0\text{–}t}$ is usually calculated by the so-called linear trapezoidal rule using the following formula:

$$AUC_{0-t} = \sum_{1}^{t} \left(\frac{C_i + C_{i-1}}{2} \right) (T_i - T_{i-1})$$

Columns four to six in **Table 1** show the calculation procedure for the anagrelide data, yielding a value of $AUC_{0\text{–}t}$ = 12.03 ng.h/mL

λ and $T_{1/2}$

For calculating the remaining PK parameters, a statistical model must be established. In general, there is no special parametric PK model that fits the 'plasma concentration over time' profile for the whole time interval. Nevertheless, it is empirically accepted that a single exponential model can be fitted to describe the 'plasma concentration over time' profile during the so-called terminal or elimination phase, and that this declining exponential curve will continue even beyond the observation interval. In other words, it is assumed that the disappearance of the drug molecules follows the most simple linear one-compartmental model (treating the body as a single compartment) during the elimination phase, but not for the entire interval. Based on the above assumption, the other PK parameters can be derived.

In most cases, elimination of the drug is a first-order process (ie, the rate of drug elimination is directly proportional to the concentration of the drug), and a log transformation makes it possible to draw a straight line through data from the elimination phase. The slope of this regression line in the elimination phase is equivalent to the elimination rate constant [1,2].

A log transformation of the concentration values in **Table 1** is plotted in **Figure 2**, from which the elimination phase and rate constant can be determined. It is crucially important to select the correct starting point for the elimination phase – ie, where elimination is no longer influenced by absorption and after which the transformed concentration values tend to be linear. This time point is the first point of the elimination phase, while the last data point measured is the last point of the elimination phase. From the anagrelide data in **Figure 2**, it is obvious that the elimination phase is from 2 to 8 hours and that the log concentration data are very close to the fitted regression line:

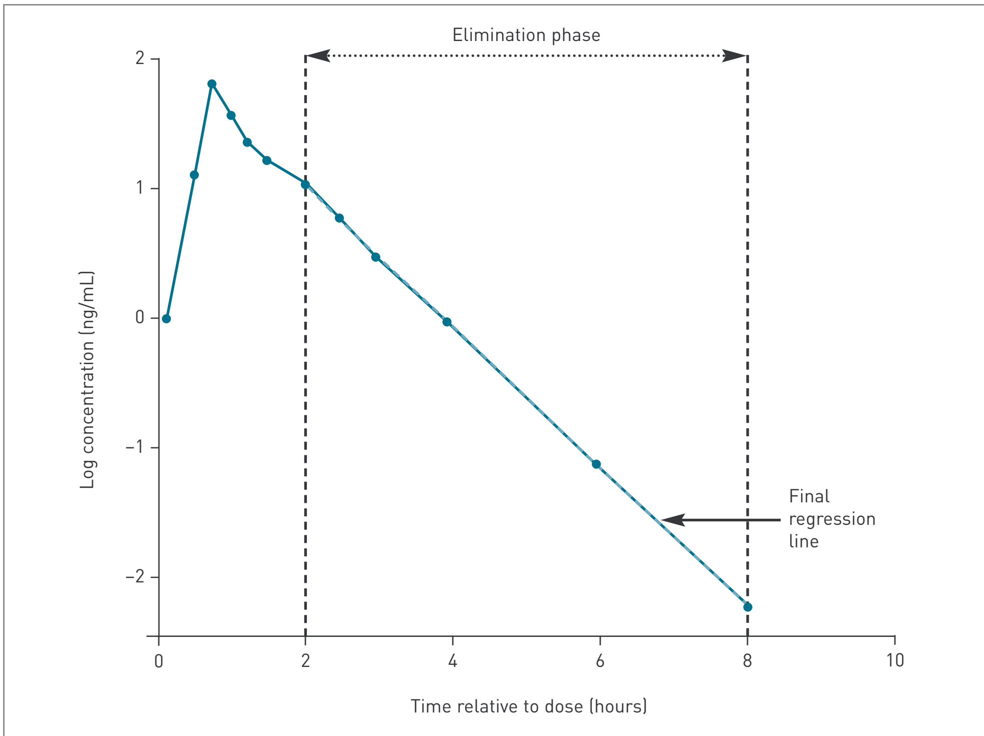
$$\text{Log}(C_i) = 2.25 - 0.58 \times T_i \text{ with } R^2 = 0.9998$$

This gives $\lambda = 0.58$ (the loss of drug activity per hour). R^2 is a measurement of the goodness-of-fit of the regression line (a value close to 1 means that the regression line is a very good fit to the data), 2.25 is the intercept, and –0.58 is the slope of the fitted regression line.

As elimination is a first-order process, by simply dividing 0.693 (ln 2) by the estimated λ value, the half-life ($T_{1/2}$) can be calculated [1,2]:

$$T_{1/2} = \ln 2 / \lambda = 0.693 / 0.58 = 1.19$$

Figure 2. Determination of the elimination phase and elimination rate constant.



$AUC_{0-\infty}$

The next step in the PK parameter calculation is to obtain the $AUC_{0-\infty}$ – the total amount of drug present in the blood – by extending the 'plasma concentration over time' profile to infinity.

Assuming that the exponential elimination process will continue beyond the last observed concentration at time *t*, the extended area after *t* is C_t / λ [1–4]. This gives:

$$AUC_{0-\infty} = AUC_{0-t} + AUC_{t-\infty} = AUC_{0-t} + C_t / \lambda$$

For the anagrelide data, we get:

$$AUC_{0-\infty} = 12.03 + 0.09 / 0.58 = 12.18$$

How to collect PK samples correctly

The success of a bioequivalence trial depends on many factors, such as:

- the standardization of study procedures
- demographic and dietary factors
- analytical work

Among these factors there are two basic study-design issues related to blood sampling that deserve special attention, since they determine whether the samples can be used to fully describe the absorption, distribution, and elimination phases of the drug.

Sampling times

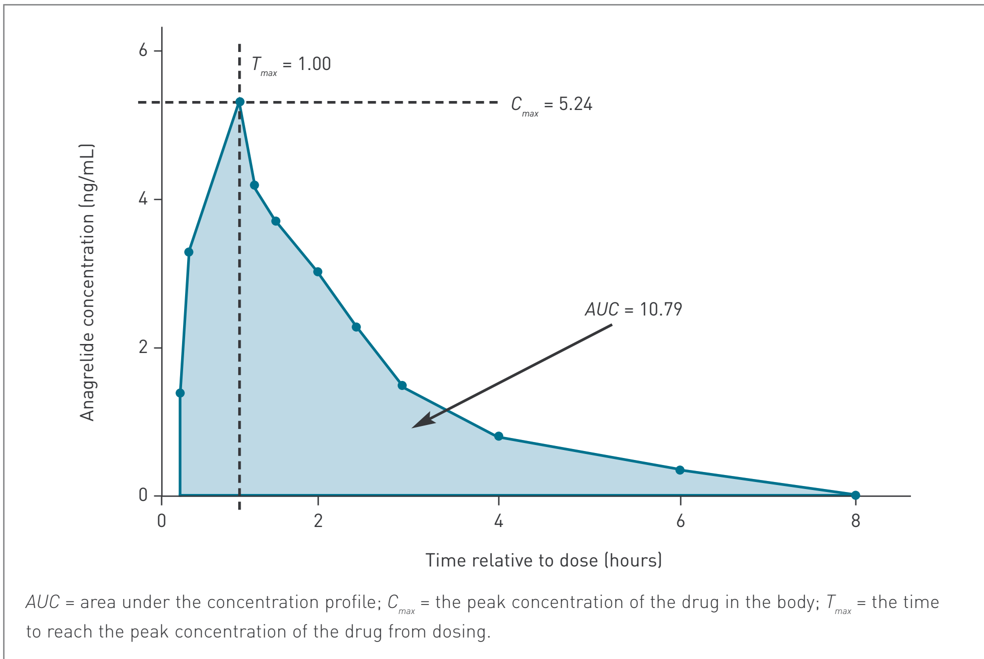
The sampling times at which the blood samples are collected have a decisive impact on the calculation of the PK parameters for the study drug. Ideally, the samples should be collected as frequently as possible during the study period so as to give an accurate PK profile. However, in practice, a relatively small number of blood samples are usually collected at selected time points due to ethical and financial considerations.

The US Food and Drug Administration (FDA) requires that sample collection should be spaced in such a way that the maximum concentration of the drug in the blood (C_{max}) and the terminal elimination rate constant (λ) can be accurately estimated [1–3]. It is important that there are enough sampling times clustered around C_{max}. For example, in **Figure 1**, the blood samples were collected every 0.25 hours from 0 to 1.5 hours around the T_{max} value (0.75 hours), meaning that the anagrelide PK profile is correct. However, occasionally, not enough blood samples are collected around C_{max} and, consequently, false C_{max}, T_{max} and AUC values are obtained. **Figure 3** shows an incorrect sampling scheme missing the time point at 0.75 hours. As a consequence of missing just this one time point, the PK parameters derived from this scheme, such as C_{max}, T_{max}, and AUC_{0-t}, are severely biased, and therefore erroneous conclusions about bioequivalence could be drawn from them.

The sampling period

The FDA requires that, to obtain an accurate estimate of λ from linear regression, sampling should continue for at least three terminal half-lives of the drug, and that at least three to four samples should be obtained during the terminal log-linear phase [1,2]. In the case of anagrelide, for example, empirical studies had shown that the half-life for this drug ranges from 1 to 2 hours, so a sampling period of 12 hours was planned for the anagrelide trial. The profiles for anagrelide from this trial show that this sampling period was long enough to obtain an accurate estimate of λ because, after 8 hours, the anagrelide plasma concentrations were no

Figure 3. A concentration profile from an incorrect sampling design.



AUC = area under the concentration profile; C_{max} = the peak concentration of the drug in the body; T_{max} = the time to reach the peak concentration of the drug from dosing.

longer quantifiable for most subjects. This is displayed for one subject in **Figure 1**. Also, the terminal phase has 6 samples, as shown in **Figure 2**, which is much greater than the required minimum of four, yielding a reliable estimate of λ for this PK profile.

What basic designs are used for bioequivalence trials?

Due to large between-subject variability in PK parameters, it is advantageous to plan bioequivalence studies with a randomized crossover design. When two drug formulations are compared, the standard two-way crossover design is often appropriate. If more than two formulations are involved in a bioequivalence study, Latin square design, which balances the period and sequence effects, becomes attractive.

A $K \times K$ Latin square design is a way of putting K replicates of each of K treatments in a $K \times K$ array such that in each row and column, all the treatments are different. A 3×3 Latin square design is shown in **Table 2** and this pattern can be extended to any size. In some crossover trials, eg, those involving nifedipine (a calcium channel blocker) and acyclovir (an antiviral drug), the PK within-

Table 2. A *3* × *3* Latin square design.

	Period 1	Period 2	Period 3
Sequence 1	A	B	C
Sequence 2	B	C	A
Sequence 3	C	A	B

subject variability (the variability of a drug's effect within a single subject) is very high. In these cases, a crossover design is no longer advantageous and a parallel design could be an alternative choice.

How do we evaluate the bioequivalence between two drugs?

Standard statistical methodology based on a null hypothesis is not an appropriate method to assess bioequivalence [4,5]. The FDA has therefore employed a testing procedure – termed the 'two one-sided tests procedure' [1–4,6] – to determine whether average values for PK parameters measured after administration of the test and reference products are equivalent. This procedure involves the calculation of a 90% confidence interval (CI) $[\theta_1, \theta_\theta]$ for the ratio (θ) between the test (T)- and reference (R)-product PK-variable averages [4,7]. The FDA guidance requires that to reach an average bioequivalence, $[\theta_1, \theta_\theta]$ must fall entirely within a range of 0.80–1.25. This is known as the bioequivalence criterion [1–3].

How do we calculate the 90% confidence interval, $[\theta_1, \theta_\theta]$?

The FDA recommends that parametric (normal-theory) methods should be used to derive a 90% CI for the quantity $\mu(T) - \mu(R)$, the mean difference in log-transformed PK parameters between the T and R products [1–3]. The anti-logs of the confidence limits obtained constitute the 90% CI $[\theta_1, \theta_\theta]$ for the ratio of the geometric means between the T and R products. The 90% CI for the difference in the means of the log-transformed data should be calculated using statistical models that are appropriate to the trial design.

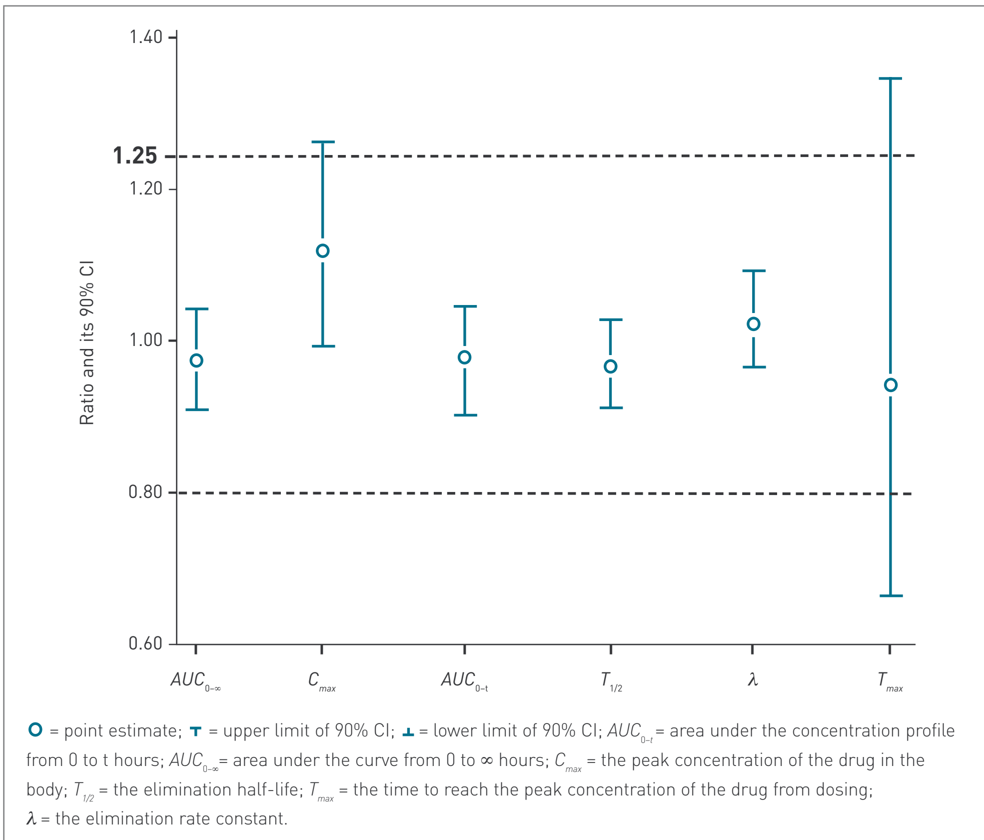
For example, for replicated crossover designs, the FDA recommends that the linear mixed-effects model (available in PROC MIXED in SAS or equivalent software [3]) should be used to obtain a point estimate and a 90% CI for the adjusted differences between the treatment means. Typically, the mixed model includes factors accounting for the following sources of variation: sequence, subjects nested in sequences, period, and treatment. The mixed model also treats the subject as a random effect so that the between-subject and within-subject variability can be measured.

Table 3. Point estimates and 90% confidence intervals for the bioavailability ratio $\mu(T) / \mu(R)$.

Parameter	Point estimate, θ	Lower 90% CI, θ_1	Upper 90% CI, θ_θ
$AUC_{0-\infty}$	0.9796	0.9151	1.0485
C_{max}	1.1237	0.9959	1.2681
AUC_{0-t}	0.9804	0.9155	1.0498
$T_{1/2}$	0.9734	0.9165	1.0338
λ	1.0273	0.9673	1.0911
T_{max}	0.9450	0.6609	1.3511

AUC_{0-t} = area under the concentration profile, ie, the amount of drug present in the blood, from 0 to *t* hours; $AUC_{0-\infty}$ = area under the curve from 0 to ∞ hours, ie, the total amount of drug present in the blood; C_{max} = the peak concentration of the drug in the body; $T_{1/2}$ = the elimination half-life; T_{max} = the time to reach the peak concentration of the drug from dosing; λ = the elimination rate constant; $\mu(R)$ = mean pharmacokinetic parameter for the reference product; $\mu(T)$ = mean pharmacokinetic parameter for the test product; θ = ratio of the geometric means between the test and reference products.

Figure 4. Ratios of pharmacokinetic parameters and their 90% confidence intervals.



○ = point estimate; ⊤ = upper limit of 90% CI; ⊥ = lower limit of 90% CI; AUC_{0-t} = area under the concentration profile from 0 to t hours; $AUC_{0-\infty}$ = area under the curve from 0 to ∞ hours; C_{max} = the peak concentration of the drug in the body; $T_{1/2}$ = the elimination half-life; T_{max} = the time to reach the peak concentration of the drug from dosing; λ = the elimination rate constant.

In **Table 3**, the ratios and 90% CIs for six PK parameters are presented as a hypothetical study of bioequivalence between test (T) and reference (R) products. These values are also displayed in **Figure 4**. For example, for $AUC_{0-\infty}$ the ratio between treatments T and R has a parametric point estimate of 0.98 and a 90% CI of 0.95–1.05. As this interval falls well within 0.80–1.25, the bioequivalence between treatments T and R can be established with respect to $AUC_{0-\infty}$. Similarly, bioequivalence holds true for AUC_{0-t}, $T_{1/2}$, and λ. However, as the 90% CIs for C_{max} and T_{max} are not completely covered by the bioequivalence acceptance range of 0.80–1.25, a conclusion of bioequivalence cannot be reached regarding the rate of absorption of the study products. Two drugs can only be considered to be bioequivalent when the rate and extent of absorption are equivalent. With respect to generic drugs it would be useful to know which statistical ranges the manufacturers adhere to.

Conclusion

We have highlighted the importance of a correctly designed bioequivalence study with respect to sampling times and sampling period. We have also demonstrated how easily biased PK parameters can be generated as a result of an inappropriate sampling scheme, leading to erroneous conclusions regarding the bioequivalence.

The typical approach to bioequivalence described in this chapter focuses on the comparison of population averages for a PK parameter. Developments in bioequivalence have included the concepts of individual and population bioequivalence that compare not only population averages, but also variance of PK parameters. Detailed discussions of these issues can be found in reference [3].

References

1. Guidance for Industry: Bioavailability and bioequivalence studies for orally administrated drug products – general considerations. Rockville, MD: US Department of Health and Human Services, Food and Drug Administration and Center for Drug Evaluation and Research, 2000.
2. Guidance for Industry: Statistical procedures for bioequivalence studies using a standard two-treatment crossover design. Rockville, MD: US Department of Health and Human Services, Food and Drug Administration and Center for Drug Evaluation and Research, 1992.
3. Guidance for Industry: Statistical approaches to establishing bioequivalence. Rockville, MD: US Department of Health and Human Services, Food and Drug Administration and Center for Drug Evaluation and Research, 2001.
4. Chow SC, Liu JP. *Design and Analysis of Bioavailability and Bioequivalence Studies*, 2nd edition. New York: Marcel Dekker, 2000.

5. Hauck WW, Anderson S. A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. *J Pharmacokinet Biopharm* 1984;**12**:83–91.
6. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinet Biopharm* 1987;**15**:657–80.
7. Westlake WJ. Statistical aspects of comparative bioavailability trials. *Biometrics* 1979;**35**:273–80.