

Web-based training program

This training session is ***LISTEN ONLY***

Background music will be playing until the beginning of the webinar

We will begin promptly at 10:00 (Eastern time)

You may submit questions to the presenter using the Question link found on the upper right hand side of the screen

Dec 10 webinar

Introduction

- Alex Dmitrienko (Eli Lilly and Company), Chair of Distance Training, Biopharmaceutical Section of ASA

Bioequivalence

- Scott Patterson (Wyeth) and Byron Jones (Pfizer)

Handouts

- Can be downloaded from BioPharmNet's web site at <http://www.biopharmnet.com/doc/doc03002-05.html>

Dr. Scott Patterson



Senior Director of Statistical
Science, Vaccines and Infectious
Disease

Wyeth Pharmaceuticals

Dr. Byron Jones



Senior Director, Statistical
Research and Consulting Center

Pfizer Global Research and
Development

BIOEQUIVALENCE

American Statistical Association Webinar, December 10, 2008

Presenters:

Scott Patterson

Wyeth Research and Development, USA

Byron Jones

Pfizer Global Research and Development, UK

Based on material in

Bioequivalence and Statistics in Clinical Pharmacology

by Scott D Patterson and Byron Jones.

Published by Chapman and Hall CRC Press.

Objectives for today

To convey the knowledge necessary to design and analyse bioequivalence trials.

To enhance the understanding of their history and place within drug development.

To review recent developments and their implications.

Contents

Drug Development, Pharmacokinetics and Definition of Bioequivalence

History and International Bioequivalence Regulations

2×2 Cross-over Designs and Average Bioequivalence

Alternative Cross-over Designs for Bioequivalence

Practical Planning for Bioequivalence Trials

Scaled ABE and Biosimilars

Questions and Discussion

Drug Development

Taking a molecule or chemical in drug form and showing regulators that:

It's safe (when used as recommended).

It works (when used as recommended).

It's manufactured to high quality standards.

Regulatory Authorities

Table 1: Regulatory Authorities

Nation	Agency
Australia	Therapeutic Goods Administration (TGA)
Canada	Therapeutic Products Directorate (TPD)
European Union	European Agency for the Evaluation of Medical Products (EMA)
China	State Drug Administration (SDA)
Japan	Ministry of Health and Welfare (MHW)
USA	Food and Drug Administration (FDA)

Pharmacokinetics

First, define some concepts:

Pharmacokinetics: 'Movements of drugs within biological systems, as affected by uptake, distribution, binding, elimination and biotransformation; particularly the rates of such movements.'
(Stedmans Medical Dictionary 26th ed.)

PK: What the body does to a drug, (as opposed to what the drug does to the body).

ADME

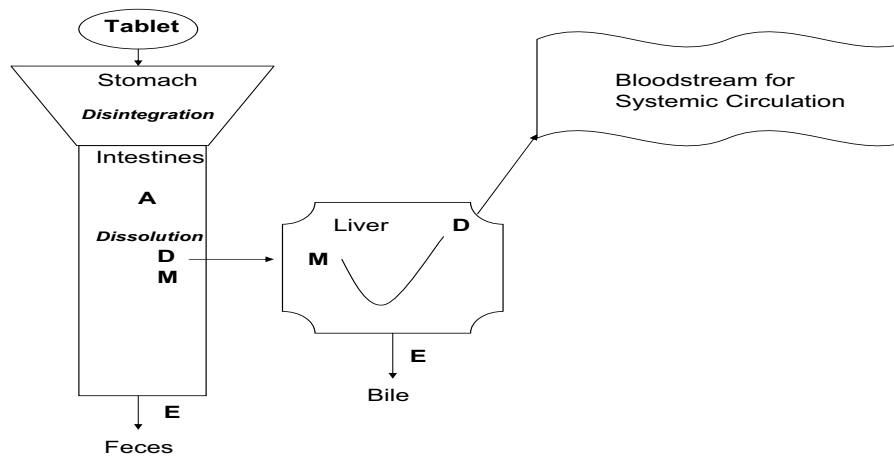


Figure 1: What happens to the drug after it's taken

7

ADME

Absorption: uptake by the body through the mouth, throat, stomach, and small/large intestine.

Distribution: how the drug substance is carried by the body through the blood to its site of action.

Metabolism: how the body breaks the drug substance into by-products.

Elimination: how the body disperses the drug product.

8

Measuring PK

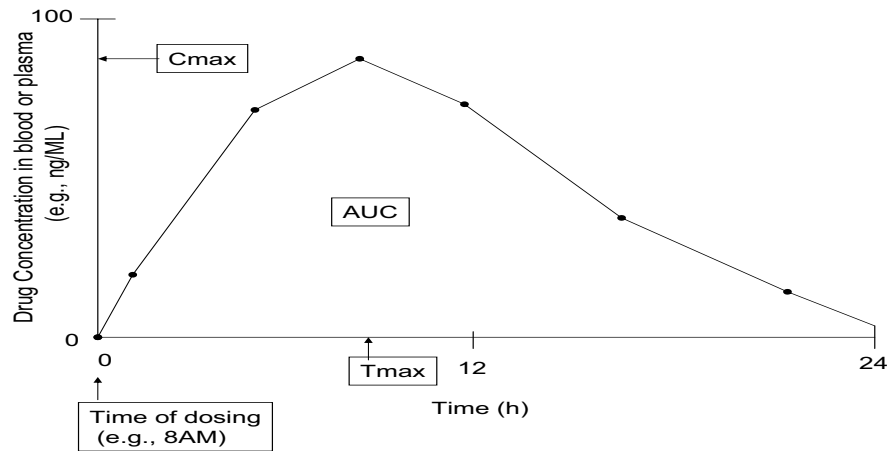


Figure 2: Plasma Concentration (ng/mL) versus Time (h)

9

Summary Measures of PK

$AUC(0-t)$ (Area under the curve from time zero to t where t is the time of last quantifiable concentration),

C_{max} (maximum concentration),

T_{max} (time of maximum concentration),

$T_{\frac{1}{2}}$ (half-life of drug substance), and

$AUC(0 - \infty) = AUC(0 - t) + \frac{C_t}{\lambda}$ where C_t is the concentration at time t and λ is -2.303 times the slope of the terminal phase of the \log_e -concentration time curve.

AUC is referred to as the extent of bioavailability.

C_{max} is the rate of bioavailability.

10

What is Bioequivalence?

Study of Pharmaceutical equivalents (same amount and type of drug) for **T**est and **R**eference formulations.

CFR Section 320.1: 'Bioequivalence means the absence of a significant difference in the rate and extent to which the active ingredient becomes available at the site of drug action.'

Area under the concentration curve (AUC, surrogate for efficacy) measures extent of exposure, and the maximum concentration (C_{max}, surrogate for safety) measures rate of exposure.

Regulators assume that formulations are not equivalent until demonstrated otherwise.

History and Regulation of BE

Objectives are to discuss:

When and how BE studies are performed

Why BE studies are performed

How to decide when formulations are bioequivalent

Potential issues with TOST bioequivalence

Current international BE regulations

‘Equivalence’

Bioequivalence (BE) studies are performed to demonstrate that different formulations or regimens of drug product are similar to each other in terms of their therapeutic benefit (efficacy) and non-therapeutic side effects (safety).

They play a key and pivotal role in the drug development process by ensuring that when a patient switches to a new formulation in the marketplace, safety and efficacy will be maintained.

Primarily, these studies are used in the study of solid oral dosage forms (i.e., drugs administered as a tablet or capsule when ingested), and we will confine discussion to this type of drug product.

Pharmaceutical Equivalence

When the new and old formulations use exactly the same substance (i.e., are pharmaceutically equivalent) why do these studies need to be done?

Rate and extent of bioavailability can be affected by very small changes in formulation.

Constituent content of the formula, small changes to the lining of the formula, and compaction into tablet (versus administration as a capsule), for example, may result in big changes in exposure.

Relative Bioavailability Studies

Many changes are made to the formulation while Phases I to II of drug development are ongoing in clinic prior to it being approved for market access.

Drug companies commonly check that these changes in formulation do not drastically change bioavailability by what are known as relative bioavailability (rel-bio) studies.

Used to ensure that the formulation to be used in Phase II or in later confirmatory trials is sufficiently similar to that used in Phase I drug development.

These are not performed to the high requirements of true bioequivalence trials.

When one wants access to the marketplace for a new formulation, a higher standard is to be met.

Prior to Marketing Approval

Bioequivalence studies are primarily used by pharmaceutical sponsors of new drug entities who have conducted pivotal confirmatory trials with a specific formulation of a drug therapy but need market access for a more commercially suitable formulation (i.e., that can be mass produced).

BE studies can be viewed as providing necessary and sufficient reassurance to regulators that the formulation to be marketed is the same as that used in the clinical confirmatory trials without the need to repeat the development programme or to perform a therapeutic equivalence study in patients with clinical endpoints.

Post-Marketing

Bioequivalence studies must also be performed following substantial post-marketing formulation alteration.

They are also used by what is termed the 'generic' pharmaceutical industry to gain market access for formulations of established drug therapies when the patent of the original sponsor's formulation expires.

When the original sponsors themselves perform a formulation change (for instance, change the site of manufacture) following approval, they often also must do a bioequivalence study to convince regulators that the new formula is safe and effective to market.

Generic Substitution

Multiple companies may produce and market similar formulations to the original marketed product following patent expiration, provided they can demonstrate bioequivalence to the original product.

Generic substitution has thus provided a means of supplying the market with inexpensive, efficacious, and safe drug products without the need to repeat an entire clinical and clinical pharmacology development package following patent expiration.

How are BE Trials Done?

Bioequivalence studies are conducted to meet documented, **legislated** regulatory standards.

Cross-over study designs (Jones and Kenward, 2003; Senn, 2003) are typically used to study bioequivalence.

Study Design

BE studies are usually conducted in male and female healthy volunteer subjects.

Each individual subject is administered two formulations (**T**est or **R**eference in the example) in one of two sequences of treatments (e.g., RT and TR).

Schematic

Table 2: Schematic Plan of a 2×2 Cross-over Study

Sequence	Period			Number of
Group	1	Wash-out	2	Subjects
1(RT)	R	—	T	$n/2$
2(TR)	T	—	R	$n/2$

R=Reference, T=Test

Wash-out, Half-life, and Carry-over

Each administration separated by a washout period appropriate to the drug under study.

This ‘wash-out’ period consists of five half-lives between administrations.

Recall that half-life is determined by looking at the elimination (after C_{max}) part of the PK concentration versus time curve and is simply the length of time it takes the body to eliminate one-half of the amount of whatever drug is in the body at any given time.

In general, if five half-lives go by, little to no drug should be left in the systemic circulation.

Other Design Features

The dose of drug substance in each formulation is pharmaceutically equivalent.

Typically the formulations are not blinded (i.e., not disguised to the patient or investigator).

Random allocation of subject to sequence is done here to ensure that time-related effects (i.e., period to period differences in blood sampling timings or laboratory handling of the samples for example) can be accounted for in the analysis and are not confounded with the estimate for the difference between formulations.

Blood samples will be collected at pre-determined, regular intervals prior to and following each dose of formulation to generate the concentration versus time curves.

23

Implications

Each subject serves as their own control (i.e., we can compare T to R on each subject).

This ensures that a precise measurement of the difference in formulations can be made.

Replication (i.e., the number of patients assigned to each sequence) is chosen to ensure that the regulatory standard for demonstrating bioequivalence can be met.

To demonstrate equivalence in plasma concentration profiles, *rate* **and** *extent* of bioavailability of the drug substance in plasma must be sufficiently similar so as to meet the regulatory standard for showing that exposure of the body to the drug substance is the same between formulations.

24

‘Just Right’

AUC and C_{max} are looked at in this situation as surrogate markers for clinical efficacy and safety, respectively.

If C_{max} increases too much with the new formulation, this could lead to unwanted side-effects. On the other hand, if it decreases too much, the drug may not be effective in treating the illness. Hence the quality of manufacturing assessment focusses on ensuring these do not change ‘too much’ in the new formulation.

The definition of ‘too much’ is quite involved.

Distributions of AUC and C_{max}

AUC and C_{max} are generally assumed to be, what is referred to as, *log-normally* distributed.

Notes on C_{max}

C_{max} is dependent on the *a priori* choice of blood sampling scheme.

It is known to be generally more variable than AUC and is sometimes problematic in the assessment of bioequivalence.

Other measures of rate of absorption have been proposed in the literature such as Direct Curve Metrics and C_{max}/AUC, and 'indirect' metrics.

However, simulation based assessment of alternatives has demonstrated such measures to be less desirable than the use of C_{max} to date.

C_{max} thus seems to be held as the 'least undesirable' measure available at present for rate of bioavailability.

History

In the late 1960s and 1970s, advances in chemical engineering increased the capability to create inexpensive copies of patented drug products (since termed 'generics').

Following patent expiration, such new formulations could potentially be marketed.

This was desirable from a governmental perspective for public health.

Need for Regulation Identified

When some pharmaceutically equivalent copies of drug products were produced, reports of therapeutic failure received a great deal of public attention in the United States.

The FDA was authorised under the 1984 Drug Price Competition and Patent Term Restoration Act to create an approval process for generic drug products.

For approval to market, the FDA decided to require a bioequivalence study for market access with prespecified decision rules for acceptability based on the data collected.

Such studies were also required for extension of patent protection for innovators seeking to maintain market exclusivity.

Two One-Sided Testing

Getting it 'just right' was designated the 'two one-sided testing procedure' (TOST).

To clarify, one hypothesises that data in the new formulation are 'too low' (H_{01}) relative to the new formulation or also that they are 'too high' (H_{02}).

If both hypotheses are rejected by the data in favour of their alternatives (H_{11} , H_{12}), then the new formulation is deemed to be bioequivalent to the reference formulation.

Two One-Sided Testing

Under this approach to inference, the usual null hypothesis was reformulated to correspond to the structure of testing the question of bioequivalence:

$$H_{01} : \mu_T - \mu_R \leq -\Delta \quad (1)$$

versus the alternative

$$H_{11} : \mu_T - \mu_R > -\Delta$$

and

$$H_{02} : \mu_T - \mu_R \geq \Delta \quad (2)$$

versus the alternative

$$H_{12} : \mu_T - \mu_R < \Delta$$

Regulatory Risk

The goalpost Δ was chosen at FDA to be equal to $\ln 1.25 = 0.2231$ (corresponding to a 20% range on the natural scale).

For each of the hypotheses H_{01} and H_{02} it was determined that the FDA wanted no more than a 5% chance of a Type 1 error.

Recall that this means that the FDA wanted no more than a 5% chance that a study would demonstrate bioequivalence when in truth, the formulations were not bioequivalent.

Mathematically, for $\mu_T - \mu_R$, the TOST corresponds to showing that a 90% *confidence interval* is contained in the interval $-\ln 1.25$ to $\ln 1.25$.

Issues with TOST

Narrow therapeutic index drugs (for which a slight change in dose or exposure can cause a large alteration in response to treatment).

Such small changes in mean test to reference rate and extent of exposure are potentially clinically meaningful in a small proportion of patients, and some have advocated special equivalence definitions for narrow therapeutic index products whereby such drugs would be held to a more strict regulatory standard (e.g., equivalence limits corresponding to a ten percent range on the \log_e -scale, 0.90 to 1.11).

Other Issues with TOST

High variability products (with within-subject standard deviations in excess of 0.30), require sample sizes in excess of 30 subjects in order to have less than a 10% to 20% chance of a Type 2 error.

Some have argued that small changes in rate and extent of exposure for such products are not clinically meaningful and have advocated allowance of a less strict regulatory standard - e.g., equivalence limits corresponding to a 30% equivalence range on the \log_e -scale, i.e. 0.70 to 1.43 on the natural scale.

A Major Issue with TOST

The concept of switchability of formulations *for the individual patient* is not addressed by the TOST (also referred to as average bioequivalence).

In the TOST, *Population* means are compared, and variation between individual subjects (or patients) is factored out of the variation used to assess the distance between population means as described above.

If individual switching is important one would need to look at *individual* bioequivalence.

Under this approach, the question, asked is 'Can I safely and effectively switch my patient from their current formulation to another?'

IBE

Consideration of these individual and population bioequivalence ideas (and sundry others) led the FDA to form a bioequivalence working group in the mid-1990s.

This body (composed of FDA representatives from clinical, scientific, and statistical disciplines) was tasked with determining whether a public health risk under the average bioequivalence approach could exist and if so to determine a method or methods to evaluate bioequivalence in a manner to protect the public health.

Partial Implementation of IBE

After considering the public comments on the preliminary draft 1997 guidance, the FDA re-issued two draft guidances on the topic of bioequivalence in August 1999 replacing the draft guidance issued in 1997).

BCS

The FDA followed up in 2000 with the introduction of the 'Biopharmaceutical Classification System'.

Orally administered drug products are categorized based upon in vitro testing into classes I, II, III, or IV.

Class I compounds, known as highly soluble and permeable in that they are quick to dissolve when ingested and are absorbed directly into the body quickly, are exempt from the requirements of demonstrating bioequivalence in a clinical study and only must demonstrate that in vitro dissolution profiles for the formulations under study are equivalent.

Under the BCS guidance, only Class II, III, and IV drugs are required to demonstrate in vivo bioequivalence before being granted market access.

Return to TOST

Following additional discussion at the 2001 Pharmaceutical Sciences Advisory Committee, the FDA provided revised final guidance which removed the potential for using population and individual bioequivalence for market access.

It is possible that in future the use of these criteria will be re-investigated if the FDA determines that there is a need for such based upon observations of the marketplace.

We will consider the statistical properties of alternative methods of assessing bioequivalence later.

International BE Regulation

TOST was formalised in the 1992 FDA Guidance and applied to both pre- and post-marketing approvals for changes in formulation.

Average bioequivalence quickly became an international standard with most nations utilising the FDA's 1992 guidance or slight modifications to the approach.

This procedure was adopted as the standard method by European and Canadian regulatory authorities subsequent to finalization of the US FDA guidance in 1992.

Japan, China, and Australia also follow this procedure (with minor changes in study design or decision rules) for the assessment of bioequivalence.

2 × 2 Cross-over Designs and Average Bioequivalence with Examples

1. The two-period two-treatment design
2. Example 1
3. Plotting the data
4. Formal analysis via a linear model
5. Testing for ABE: Confidence interval approach
6. Sample size calculations

41

Two-Period Cross-over Design

Sequence	Period	
	1	2
1	R	T
2	T	R

Assume:

n_1 subjects on sequence 1 (RT)

n_2 subjects on sequence 2 (TR)

and a wash-out of at least 5 half-lives between periods 1 and 2

42

Example 1: Data

Subject	Sequence RT			
	AUC Period		Cmax Period	
	1	2	1	2
1	2849	2230	499	436
4	2790	2864	733	416
.
.
34	1737	1655	425	319
36	2040	2199	464	384

Subject	Sequence TR			
	1	2	1	2
2	2025	2000	438	361
3	2090	1826	535	558
.
.
30	2519	1941	537	400
35	1560	1629	463	372

43

Plotting the data

Before we formally analyse BE data we should plot the data in various ways.

This will alert us to any unusual features and help us get a good idea of what the formal analysis will report.

In the following we will see a number of useful plots.

44

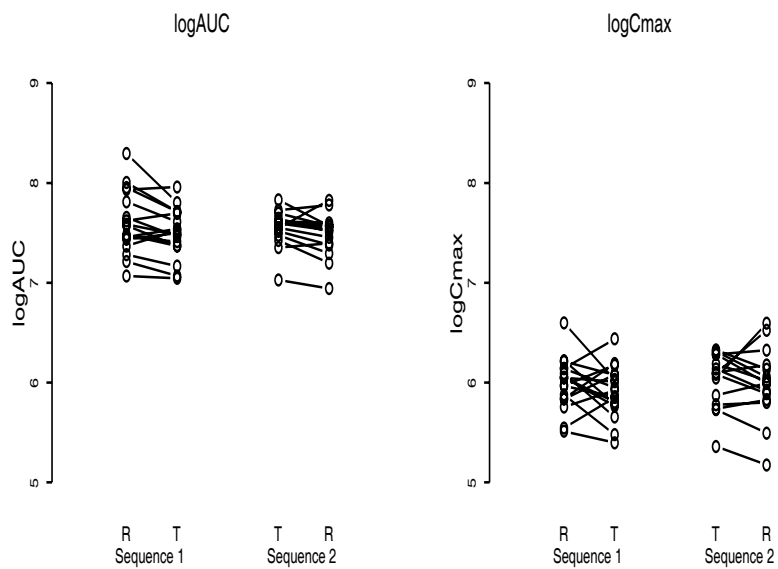
Subject-profiles plot

The subject-profiles plot is constructed for each sequence by

1. plotting on the vertical axis, for each subject, the Period 1 and Period 2 responses against the values 1 and 2, respectively, on the horizontal axis, and then
2. joining the two responses for each subject with a line.

45

Example 1: Subject profiles plot.



46

Formal Analysis of the Data

In order to produce a more formal analysis we need to specify a model for the observed data.

This linear model needs to account for those features that:

- vary systematically between the two responses on a subject
(the fixed effects)
- vary randomly between the two responses on a subject and between responses on different subjects
(the random effects).

Linear Model for Data

- Let y_{ijk} denote the (log-transformed) response obtained from subject k in period j in sequence group i ; $i = 1, 2$; $j = 1, 2$; $k = 1, 2, \dots, n_i$.
- Model will contain:
 - fixed effects for sequences, periods and formulations
 - random effects for subjects and within-subject variability.

Linear Model for Data - cont'd

Table 3: **Fixed effects**

Group	Period 1	Period 2
1(RT)	$\gamma_1 + \mu_R + \pi_1$	$\gamma_1 + \mu_T + \pi_2$
2(TR)	$\gamma_2 + \mu_T + \pi_1$	$\gamma_2 + \mu_R + \pi_2$

Note: we assume that there are no carry-over effects of formulation from Period 1 to Period 2

Linear Model for Data - cont'd

Table 4: **Random effects**

Group	Period 1	Period 2
1(RT)	$\xi_{k(1)} + \varepsilon_{11k}$	$\xi_{k(1)} + \varepsilon_{12k}$
2(TR)	$\xi_{k'(2)} + \varepsilon_{21k'}$	$\xi_{k'(2)} + \varepsilon_{22k'}$

Random effect for subject k in Group 1 and subject k' in Group 2.

Linear Model for Data - cont'd

Putting this all together we have:

$$y_{ijk} = \gamma_i + \mu_{d[i,j]} + \pi_j + \xi_{k(i)} + \varepsilon_{ijk},$$

where

- γ_i , π_j , and $\mu_{d[i,j]}$ are fixed effects for sequence, period, and formulation,
- $d[1, 1] = R$, $d[1, 2] = T$,
 $d[2, 1] = T$ and $d[2, 2] = R$,
- $\xi_{k(i)}$ and ε_{ijk} are random effects for subjects and within-subjects, respectively,
- $\text{Var}(\xi_{k(i)}) = \sigma_B^2$, the between-subject variance,
- $\text{Var}(\varepsilon_{ijk}) = \sigma_W^2$, the within-subject variance.

Estimation of $\mu_T - \mu_R$

Let

$$\bar{y}_{ij.} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk}$$

denote the mean response of the subjects in period j in sequence group i .

Table 5: **Group-by-Period Means**

Group	Period 1	Period 2
1(RT)	$\bar{y}_{11.}$	$\bar{y}_{12.}$
2(TR)	$\bar{y}_{21.}$	$\bar{y}_{22.}$

Estimation of $\mu_T - \mu_R$, cont'd

For Group 1: $E(\bar{y}_{11.} - \bar{y}_{12.}) = \pi_1 - \pi_2 + \mu_R - \mu_T$.

For Group 2: $E(\bar{y}_{21.} - \bar{y}_{22.}) = \pi_1 - \pi_2 + \mu_T - \mu_R$.

Hence

$$E\left\{\frac{1}{2}[(\bar{y}_{21.} - \bar{y}_{22.}) - (\bar{y}_{11.} - \bar{y}_{12.})]\right\} = \mu_T - \mu_R.$$

Estimation of $\mu_T - \mu_R$, cont'd

That is,

$$\hat{\mu}_T - \hat{\mu}_R = \frac{1}{2}(\bar{y}_{21.} - \bar{y}_{22.} - \bar{y}_{11.} + \bar{y}_{12.})$$

and

$$\text{Var}(\hat{\mu}_T - \hat{\mu}_R) = \frac{1}{4} \left[\frac{\sigma_W^2}{n_1} + \frac{\sigma_W^2}{n_1} + \frac{\sigma_W^2}{n_2} + \frac{\sigma_W^2}{n_2} \right] = \frac{\sigma_W^2}{2} \left[\frac{1}{n_1} + \frac{1}{n_2} \right].$$

If $n_1 = n_2 = n/2$, then

$$\text{Var}(\hat{\mu}_T - \hat{\mu}_R) = \frac{\sigma_W^2}{2} \left[\frac{2}{n} + \frac{2}{n} \right] = \frac{2\sigma_W^2}{n}.$$

Estimation of $\mu_T - \mu_R$, cont'd

Table 6: **Example 1: Group-by-Period Means**

logAUC			
Group	Period 1	Period 2	Mean
1(RT)	$\bar{y}_{11.} = 7.60(17)$	$\bar{y}_{12.} = 7.50(17)$	$\bar{y}_{1..} = 7.55$
2(TR)	$\bar{y}_{21.} = 7.55(15)$	$\bar{y}_{22.} = 7.48(15)$	$\bar{y}_{2..} = 7.51$
Mean	$\bar{y}_{.1.} = 7.58$	$\bar{y}_{.2.} = 7.49$	$\bar{y}_{...} = 7.53$

logCmax			
Group	Period 1	Period 2	Mean
1(RT)	$\bar{y}_{11.} = 5.99(17)$	$\bar{y}_{12.} = 5.91(17)$	$\bar{y}_{1..} = 5.95$
2(TR)	$\bar{y}_{21.} = 6.02(15)$	$\bar{y}_{22.} = 5.99(15)$	$\bar{y}_{2..} = 6.01$
Mean	$\bar{y}_{.1.} = 6.01$	$\bar{y}_{.2.} = 5.95$	$\bar{y}_{...} = 5.98$

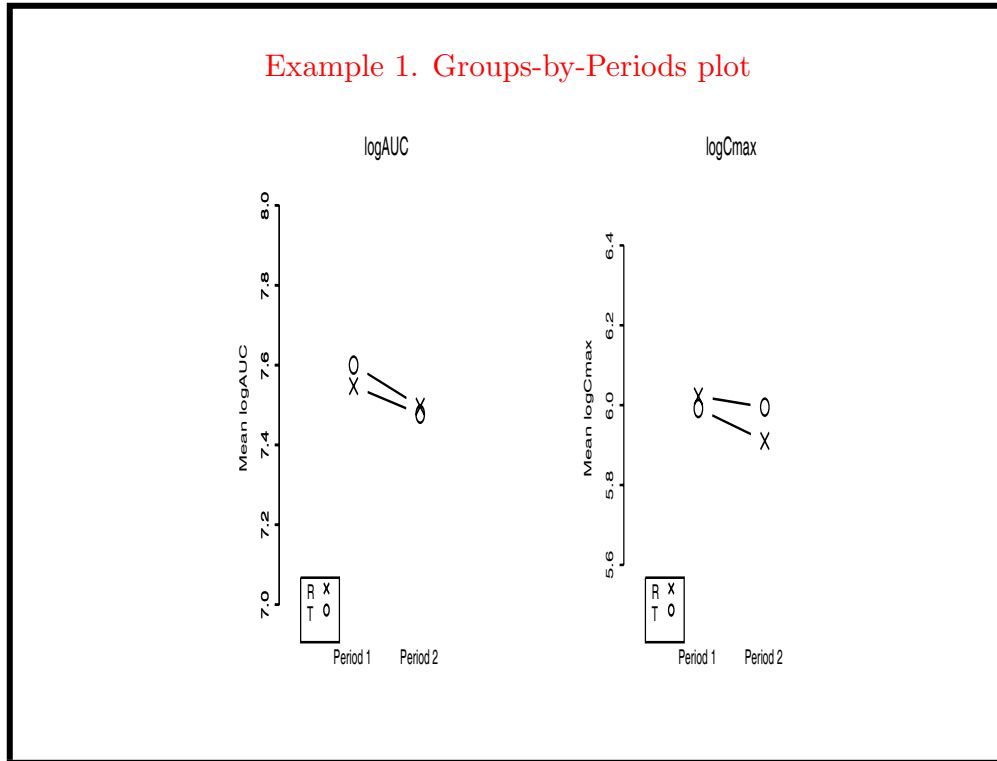
(sample size in brackets)

Groups-by-Periods Plot

The difference between the means of T and R over the two periods in each group can be displayed in the groups-by-periods plot.

Here, the four mean $\bar{y}_{11.}$, $\bar{y}_{12.}$, $\bar{y}_{21.}$ and $\bar{y}_{22.}$ are plotted against their respective period labels and the two means from the same group are joined up.

Example 1. Groups-by-Periods plot



57

Estimation of $\mu_T - \mu_R$, cont'd

Log AUC:

$$\hat{\sigma}_B^2 = 0.0516, \hat{\sigma}_W^2 = 0.0110$$

$$\hat{\mu}_T - \hat{\mu}_R = -0.0166, \text{ Std error} = 0.0267.$$

Log Cmax:

$$\hat{\sigma}_B^2 = 0.0453, \hat{\sigma}_W^2 = 0.0383$$

$$\hat{\mu}_T - \hat{\mu}_R = -0.0269, \text{ Std error} = 0.0490.$$

58

Testing for Average Bioequivalence (ABE)

ABE is demonstrated if

the 90% two-sided confidence interval for $\mu_T - \mu_R$

falls within the acceptance limits of

$\ln(8/10) = -0.2231$ and $\ln(10/8) = 0.2231$.

$\ln(10/8)$ is a goalpost, set by regulatory agencies that define how 'close' the two formulations must be to be declared bioequivalent.

When exponentiated the limits are 0.80 and 1.25.

59

Confidence Interval for $\mu_T - \mu_R$

If $\hat{\sigma}_W^2$ is an estimate of σ_W^2 on $n - 2$ degrees of freedom (d.f.) and $t_{0.95}(n - 2)$ is the upper 95% percentile of the t -distribution on $n - 2$ d.f., the 90% confidence interval for $\mu_T - \mu_R$ is

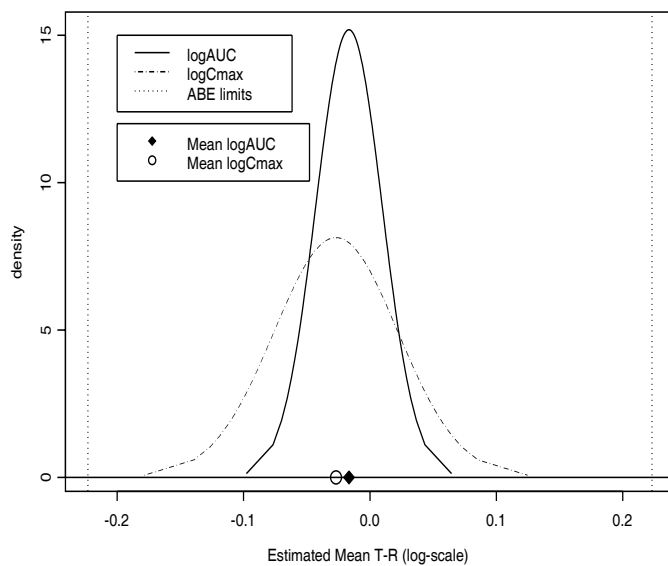
$$\hat{\mu}_T - \hat{\mu}_R \pm t_{0.95}(n - 2) \sqrt{\frac{\hat{\sigma}_W^2}{2} \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}.$$

60

Table 7: **Example 1: Confidence Intervals**

Endpoint	$\hat{\mu}_T - \hat{\mu}_R$	90% Confidence Interval
logAUC	-0.0166	(-0.0612, 0.0280)
logCmax	-0.0269	(-0.1102, 0.0563)
Endpoint	$\exp(\hat{\mu}_T - \hat{\mu}_R)$	90% Confidence Interval
AUC	0.98	(0.94, 1.03)
Cmax	0.97	(0.90, 1.06)

Example 1: Fitted normal densities for $\hat{\mu}_T - \hat{\mu}_R$.



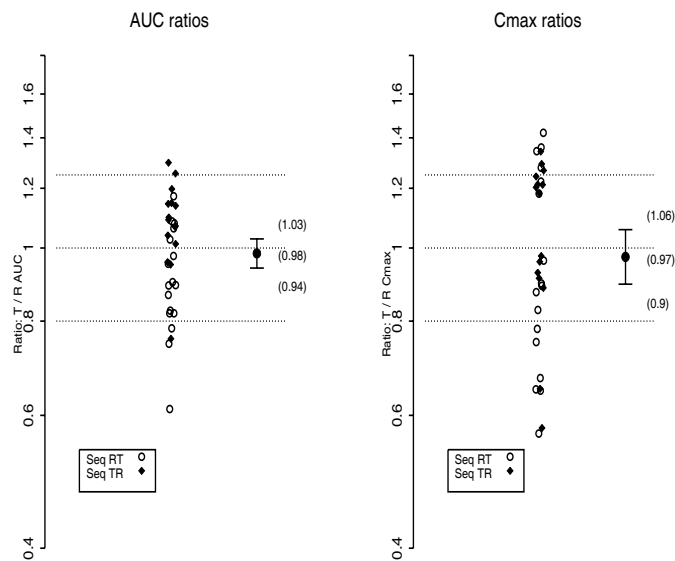
Final Summary for Example 1.

We display the confidence intervals on the natural scale alongside a plot of the ratios T:R for each of AUC and Cmax in the following figure. We note that, for Cmax especially, there are many subjects that have ratios outside the ABE limits of (0.8, 1.25).

This example highlights that fact that to be equivalent on the ABE criterion it is only necessary to show that the means of T and R do not differ to a significant extent.

63

Example 1: 90% Confidence intervals for $\exp(\hat{\mu}_T - \hat{\mu}_R)$.



64

The Two One-sided Testing Procedure (TOST)

The confidence interval approach to testing for ABE directly corresponds to an hypothesis testing procedure suggested by Schuirmann (1987):

To determine that T and R are ABE we must reject **both** of the following one-sided hypotheses:

$$H_{01} : \mu_T - \mu_R \leq -\ln 1.25$$

or

$$H_{02} : \mu_T - \mu_R \geq \ln 1.25.$$

If both of these are rejected then we can conclude that

$$-\ln 1.25 < \mu_T - \mu_R < \ln 1.25.$$

TOST for Example 1

Example 1

LogAUC:

t-statistic for $H_{01} = 7.86$ on 30 d.f. (P-value ≈ 0);
for $H_{02} = -9.12$ on 30 d.f. (P-value ≈ 0).

LogCmax:

t-statistic for $H_{01} = 4.00$ on 30 d.f. (P-value = 0.0002);
for $H_{02} = -5.12$ on 30 d.f. (P-value ≈ 0)

Sample size calculations

Methods for calculating the sample size based on the TOST procedure are described in Patterson and Jones (2005).

The values of δ and σ_W have been expressed relative to μ_R , i.e., $\delta = 100(\mu_T - \mu_R)/\mu_R$ and σ_W is a fraction of μ_R ($f \times \mu_R$). So, for example, a value of 0.2 implies $\sigma_W = 0.2\mu_R$.

Table 8: Samples sizes for a 2×2 cross-over trial to detect ABE

$f(\text{for } \sigma_W)$	δ	90% Power
	0	20 (91.92)
0.2	5	26 (91.64)
	10	48 (90.77)
	15	130 (90.11)

TRIALS WITH FOUR PERIODS

CONTENTS

1. Four-period, two-treatment design
2. Example 4

Four-period Designs

1.	R	R	T	T	2.	R	T	R	T	3.	R	T	T	R
	T	T	R	R		T	R	T	R		T	R	R	T
4.	R	T	R	R	5.	R	R	T	R	6.	R	T	T	T
	T	R	T	T		T	T	R	T		T	R	R	R
					7.	R	R	R	T					
						T	T	T	R					

Table 9: Efficiencies of Designs 1 through 7

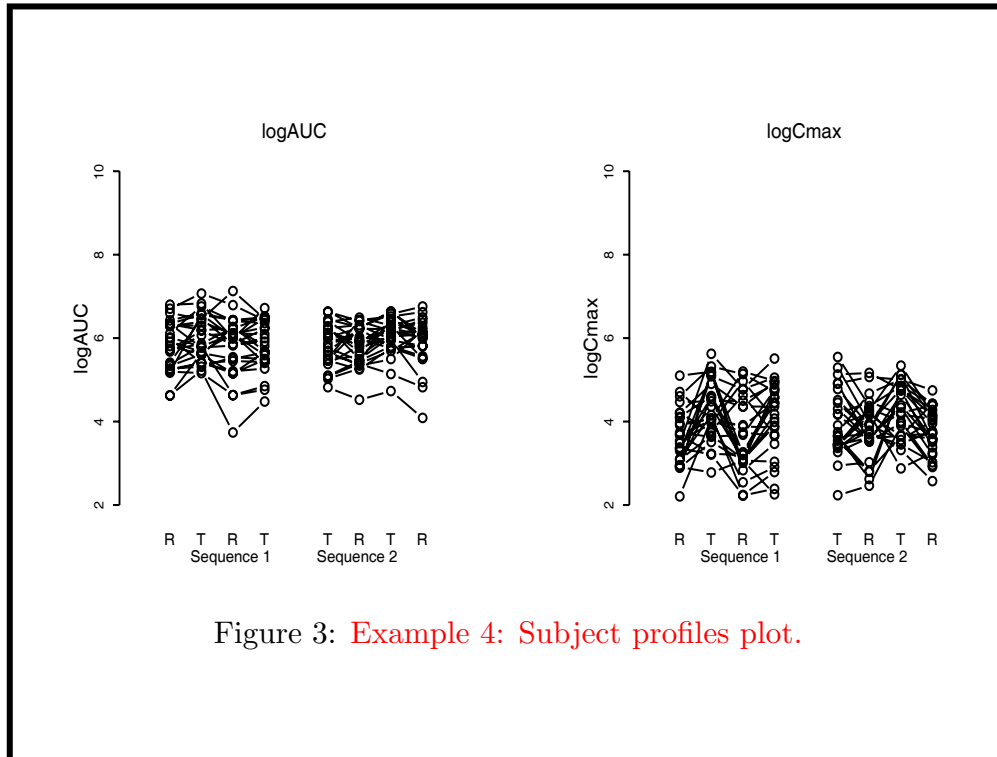
Design	Adjusted for carry-over	Unadjusted for carry-over
1	90.91	100.00
2	18.18	100.00
3	90.91	100.00
4	54.55	75.00
5	54.55	75.00
6	72.73	75.00
7	66.67	75.00

Example 4: Data

Sequence RTRT								
	AUC Period				Cmax Period			
Subject	1	2	3	4	1	2	3	4
1	812.60	1173.70	889.10	620.10	99.85	204.09	170.94	112.78
3	545.10	542.90	-	-	67.69	41.73	-	-
.
54	185.20	222.90	182.10	194.10	18.34	16.09	21.50	9.57
57	180.60	174.70	102.90	117.00	9.10	58.44	12.74	18.33

Sequence TRTR								
	AUC Period				Cmax Period			
Subject	1	2	3	4	1	2	3	4
2	216.30	338.00	502.80	398.60	29.06	50.48	35.15	55.71
4	632.60	520.00	716.70	860.40	91.25	43.86	168.78	61.04
.
55	246.90	620.90	678.30	752.20	42.20	106.69	150.52	115.15
56	235.40	190.40	318.30	248.40	39.15	13.79	122.03	62.32

71



72

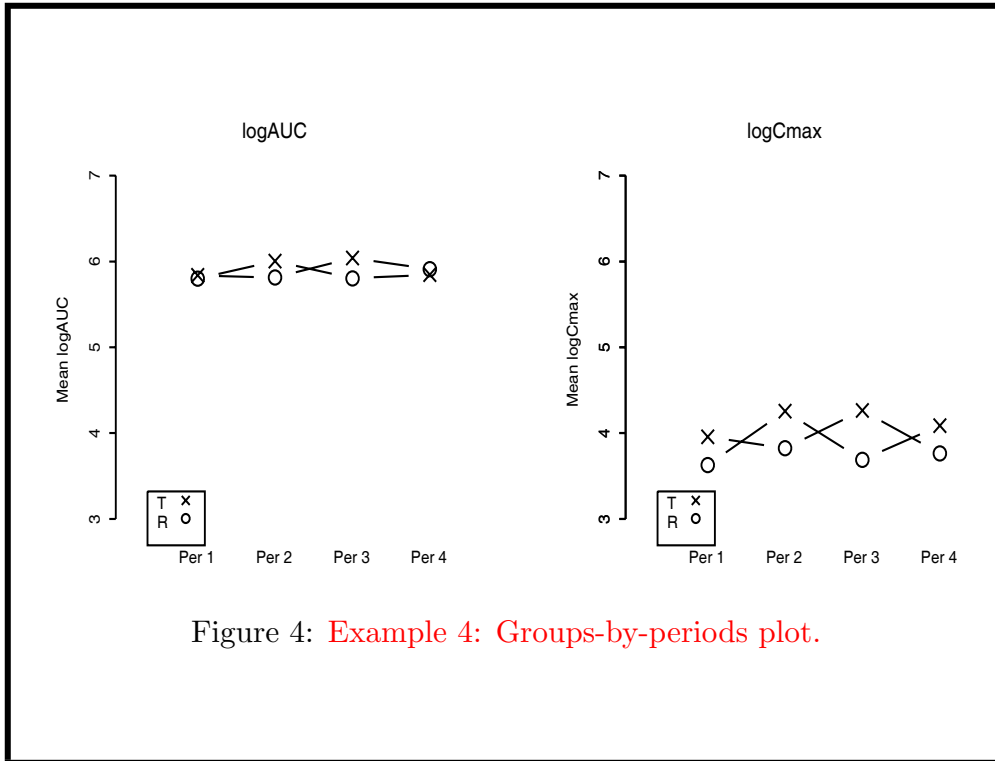


Figure 4: Example 4: Groups-by-periods plot.

73

Table 10: Example 4: TOST Procedure Results

Endpoint	$\hat{\mu}_T - \hat{\mu}_R$	90% Confidence Interval
logAUC	-0.1719	(-0.2630, -0.0809)
logCmax	-0.1299	(-0.2271, -0.0327)
Endpoint	$\exp(\hat{\mu}_T - \hat{\mu}_R)$	90% Confidence Interval
AUC	0.84	(0.77, 0.92)
Cmax	0.88	(0.80, 0.97)

Evidence for lack of ABE on both metrics.

74

Practical Planning for BE Trials

Failing BE

Carry-over

Outliers

Algorithm: Planning a Bioequivalence Study

1. Determine the number of formulations (and doses) to be studied for bioequivalence.
2. Calculate the sample size for a standard cross-over (i.e. a non-replicate 2×2 , three or four-period) design.
3. Consider available clinical resources.
4. For products with low to moderate intra-subject variation ($CV_W < 30\%$) where adequate resources are available, use the standard cross-over design.
5. For highly-variable products, where sample size exceeds available resources, consider a replicate cross-over design and re-assess sample size (50% of standard).

Failing BE

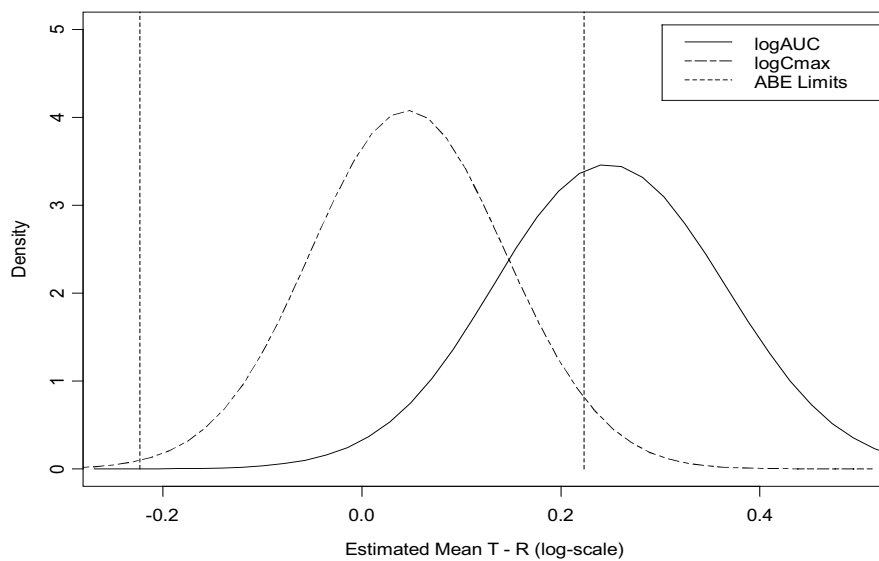
For some drug products, even if one tries time and time again to demonstrate bioequivalence, it may be that it just can not be done.

A common misconception is that this means that the test and reference formulations are bioinequivalent.

This is not necessarily the case.

77

Potentially *Bioinequivalent* Products: Fitted Normal Densities



78

Failing BE

Formulations may fail to show bioequivalence for several reasons:

1. the estimated $\mu_T - \mu_R$ lies too far from zero,
2. variation is greater than expected resulting in too wide a confidence interval for $\mu_T - \mu_R$,
3. insufficient sample size is used (also yielding too wide a confidence interval for $\mu_T - \mu_R$),
4. or some combination of these.

Failing BE

Failure to demonstrate bioequivalence is therefore different but related to bioinequivalence.

Only in cases where sample size is very large and point estimates for δ lie outside the acceptance bounds would one definitely conclude bioinequivalence.

Next Steps when Failing BE

Not always necessary in regulatory science to secure approval of a new product.

Task then is to model exposure's (AUC, C_{max}) relationship to efficacy and safety in patients using the reference formulation's clinical data.

If therapeutic equivalence can be shown from such an exercise, then approval may be obtainable.

In the knowledge of the extent to which the test formulation changes exposure (measured in the bioequivalence study or studies), one then *simulates* what a change of the magnitude observed for AUC and C_{max} for the test formulation would result in in terms of patient response in clinical use.

Failing BE

Simulation-based dose-exposure-response assessment is limited in scope of application *to only new (i.e. innovative) products*.

Existing marketed products may not apply such a procedure and must demonstrate average bioequivalence to have access to market (in most cases).

Carry-over

This refers to the occurrence of a non-zero plasma concentration of drug in a sample prior to dosing.

As such it complicates the analysis of bioequivalence data, by aliasing or biasing the assessment of changes between formulations.

To prevent this, a wash-out period (of at least five half-lives) is employed to prevent such occurrences between each study period.

Carry-over

Carry-over is very unusual but not unknown in bioequivalence studies and can arise from a variety of factors. Some are:

1. Long-half life drugs (with inadequate, too short, wash-out duration),
2. Serendipitous inclusion in the trial of subjects who poorly metabolise or eliminate the drug, and
3. Random occurrences (possibly due to assay problems).

Carry-over

As a practical matter, even if a more than adequate wash-out is used, there will be instances where pre-dose concentrations in periods after the first are non-null.

FDA guidance (2003) recommends that:

If the predose concentration is less than or equal to 5 percent of C_{max} value in that subject, the subjects data without any adjustments can be included in all pharmacokinetic measurements and calculations. If the predose value is greater than 5 percent of C_{max} , the subject should be dropped from all BE study evaluations.

Carry-over

This guidance tacitly assumes that the occurrence of carry-over of sufficient magnitude to impact inference is random in line with recent publications on the topic (Jones and Kenward, 2003).

This approach has the benefit of simplicity, and given the sparsity of the occurrence of relevant carry-over, it is expected that its application will suit most circumstances.

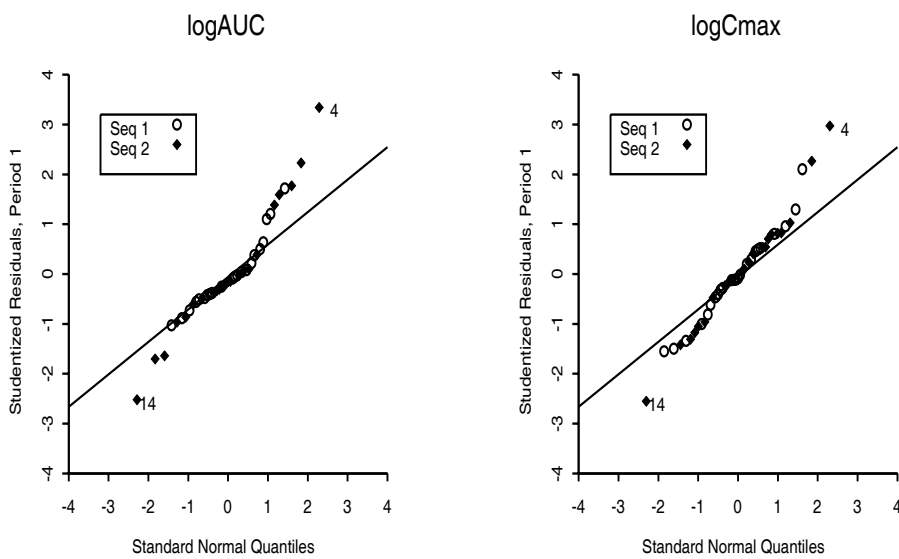
Outliers

Although not explicitly stated in regulatory bioequivalence guidance, there is a very great distinction between an outlier in a statistical sense and in a regulatory sense.

In statistical training, outliers are generally introduced as a topic related to assessment of model fit.

An outlier is defined as a residual (observed less predicted) data point that has large value - i.e. the model does not fit the data point well.

Outlier Example



Outliers

Outliers occur frequently in bioequivalence trials simply by chance.

However, outlier detection is held to be potentially indicative of product failure or subgroup identification and is scrutinised carefully by regulatory authorities prior to approval.

Statistical detection of an outlier **is not** sufficient reason to exclude a subject's observation(s).

If data were to be excluded 'scientific evidence or explanations' (FDA, 1992) should be supplied.

An example of an acceptable reason might be if it could be documented that a subject failed to swallow their medication or took too much.

Outliers

In terms of statistical impact, an outlier (or set of outliers) may impact the estimate $\hat{\delta}$ (by influencing its position relative to 0) and inflate the estimate of within-subject variance $\hat{\sigma}_W^2$ (resulting in a wider than expected confidence interval) or both.

Impacting either of these parameters implicitly makes it more difficult to demonstrate average bioequivalence.

Outliers

Whether an outlier is a product failure, a subject-by-formulation interaction, or a random event is immaterial.

These are generally confounded, and final inference with regard to bioequivalence (and regulatory approval) is based on the full data set (i.e., including the outliers).

On a practical level, this essentially means in practice that there is no such thing as an outlier in a bioequivalence data set, and while we recommend that statisticians always check the assumptions of their model, in this context there is little utility in spending too much time worrying about outliers' impact on the findings.

Preventing Outlier Induced BE Failures

As the impact of outliers cannot be controlled after the study completes, the best way to deal with them is to acknowledge that they can happen at random and to protect the study's power for random appearance of outliers at the design stage.

To do so, it is recommended that bioequivalence studies be powered at 90% and that such trials have at least 80% power under potential inflation of the variability estimate and for potential changes in δ of up to 5%.

Average Bioequivalence Deficiencies

Concerned only with the means of the distributions.

Does not take account of the variability of the distribution.

No assessment of individual subjects.

Highly variable drug products are problematic; sample size becomes large (though this can somewhat be mitigated by the use of partial replicate and replicate designs).

Regulatory limits may be too wide for low variability drugs?

ABE - Has it failed the market?

No documented evidence of therapeutic failure under 1992 ABE Guidance - See Patterson (2001) A Review of the Development of Biostatistical Design and Analysis Techniques for Assessing *In Vivo* Bioequivalence: Part One and Two. *Indian J Pharm Sci*, **63**, 81-100; 169-186.

Alternatives do not correct deficiencies and may cause more problems - see Barrett, Batra, Chow, Cook, Gould, Heller, Lo, Patterson, Smith, Stritar, Vega, Zariffa. (2000) PhRMA Perspective on Population and Individual Bioequivalence. *J Clin Pharmacol*, **40**, 561-572.

IBE and PBE were proposed stimulating debate....

Expectations for BE in Practice

To call something equivalent implies a context or criteria for the determination. There are several stakeholders in determining such a criteria:

Statistical considerations: the approach should be quantifiable, accurate, precise, well understood, and should be transparent in interpretation.

Sponsor considerations: Using a well-designed, controlled, and reasonably sized study (or set of studies) the sponsor should be able to show the criteria have been met with a quantified chance of success.

Regulatory and public-health considerations: The approach used must protect public health (in that the risk of false positive market access must be controlled at a pre-determined rate).

Further and New Developments

Highly variable drugs – Scaled Average Bioequivalence.

Biosimilar Drugs

Bioequivalence Approaches for Highly Variable Drugs

FDA investigated various approaches (Haidar et al., 2008) that would reduce the sample size without allowing inequivalent products to reach the market.

High variability implies the coefficient of variation is 30% or greater.

Four different approaches were considered and compared using simulation.

One method recommended.

97

Four Approaches for Highly Variable Drugs

1. Direct expansion of the bioequivalence limits.
2. Expansion of the bioequivalence limits based on a fixed sample size.
3. Widening the bioequivalence limits based on the Reference variability.
4. Expansion of the bioequivalence limits based on sample size and variability.

Recommended widening the bioequivalence limits based on the Reference variability.

98

Widening the Bioequivalence Limits Based on the Reference Variability.

This is the Scaled Average Bioequivalence approach:

$$\frac{(\mu_T - \mu_R)^2}{\sigma_{WR}^2} \leq \theta.$$

Recommended design is three-period cross-over with sequence
TRR, RTR and RRT.

$$\theta = \frac{(\ln \Delta)^2}{\sigma_{W0}^2},$$

where $\Delta = 1.25$ and $\sigma_{W0} = 0.25$.

Biosimilars

Source of information: Schellekens, H. (2008) and EMEA
(European) CHMP Guideline 2005.

Definition: drugs produced using a living system or genetically
modified organism.

Differ from traditional medicines as their molecules are much
larger.

Are produced by living cells, which exhibit heterogeneity.

Biosimilars

Manufacture is complicated, requires careful monitoring using in-house standards.

Analytical methods need to be adapted for each specific product.

Existing processes will have undergone continuous improvement based on increasing experience.

In short - it is very difficult for a biologic to be replicated by a generic manufacturer.

Regulatory Situation

European Guidance has been Issued (EMEA CHMP/437/04).

In the EU, two biosimilar growth hormones have been approved, an IFN α -2b has been rejected.

In the USA, the FDA has to wait for legislation on “follow-on-biologics”.

But recombinant DNA -derived insulins and growth hormones admitted under ordinary Drug Act before specific legislation for biosimilars was introduced.

Regulatory Situation

Expected that US Congress will introduce legislation in 2009.

Experts expect large differences between US and EU legislation.

WHO considering formulation of basic guidelines for countries with less well-developed regulatory systems.

Biosimilars: Where can statisticians contribute?

Many sources of variability to explain.

Not all can be studied in a clinical trial, e.g., manufacturing process.

Additional References

CHMP (2005) Guideline in Similar Biological Medicinal Products. CHMP/437/04. European Medicines Agency.

Haidar, S.H. et al. (2007) Bioequivalence approaches for highly variable drugs and drug products. *Pharmaceutical Research*, **25**, 237–241.

Schellekens, H. (2008) The first biosimilar epoetin: but how similar is it? *Clinical Journal American Society of Nephrology*, **3**, 174–178.

Objectives for today

To convey the knowledge necessary to design and analyse bioequivalence trials.

To enhance the understanding of their history and place within drug development.

To review recent developments and their implications.