Sankhyā : The Indian Journal of Statistics
Special Issue on Biostatistics
2000, Volume 62, Series B, Pt. 1, pp. 149–161

PROFILE SIMILARITY IN BIOEQUIVALENCE TRIALS

By DAVID T. MAUGER
and
VERNON M. CHINCHILLI
The Pennsylvania State University, Hershey

SUMMARY. In a typical bioequivalence trial, summary measures of the plasma concentration versus time profile are used to compare two formulations of a drug product. Commonly used measures include: area under the curve (AUC), maximum plasma concentration (C_{max}) and time to maximum concentration (T_{max}). Equivalence of these summary measures, in general, does not guarantee equivalence of the entire profile. Rescigno (1992) and Chinchilli and Elswick (1997) propose summary statistics which measure profile similarity, but are not easily interpreted pharmacologically. We propose a method for assessing bioequivalence over the entire profile which has a familiar interpretation and represents a compromise between the insensitivity to pattern differences of summary measures and the oversensitivity of pointwise comparisons.

1. **Introduction**

The purpose of a bioequivalence trial is to determine whether or not two different formulations of a drug have similar bioavailabilities. The term *bioavailability* typically is used as a general reference to the rate at and extent to which the active drug ingredient is absorbed from a drug product. *Similarity* is usually assessed in terms of the relative difference between the two formulations. In a typical bioequivalence trial one of the drug products is a reference formulation and the other a test formulation. Each subject is administered both formulations in a randomized two-period crossover fashion. In a single-dose study blood plasma samples are taken serially, over a period of hours, and the drug concentration in each sample is estimated by assay. The curve resulting from a line plot of drug concentration versus sampling time is called the plasma drug concentration profile.

Figure 1 shows the plasma drug concentration profiles for 26 subjects in a bioequivalence trial of ibuprofen taken from Chinchilli and Elswick (1997). A standard 2×2 crossover design was used with a one-week washout period between treatment periods. Blood samples were taken at 14 time points over the first 16 hours after the formulation was administered. The profiles from Figure 1 corresponding to subjects number 9, 17 and 26 will be used illustratively throughout this paper. The first two

AMS (1991) *subject classification.* 92B15, 62P10.
Key words and phrases. Bioequivalence, bioavailability, AUC.

appear to reflect differences in the absorption rates of the two formulations while the third serves as a control in which the profiles appear fairly similar.

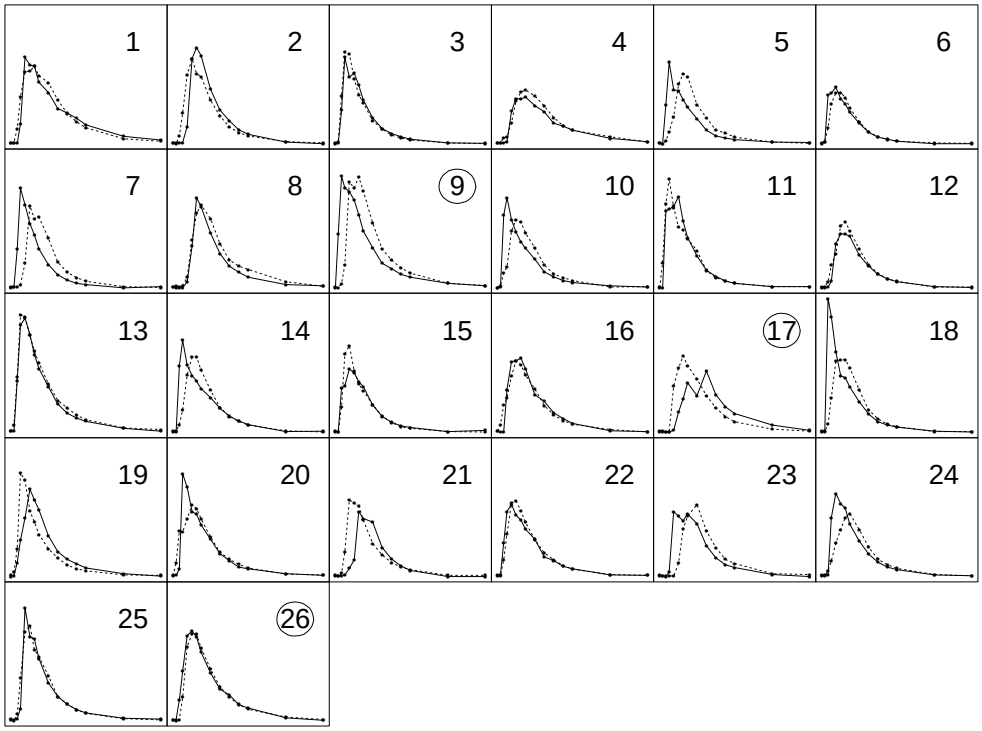

Figure 1: Plasma drug concentration profiles of two drug formulations in twenty-six subjects; solid line denotes reference formulation and dashed line denotes test formulation.

Various summary measures of this profile have been proposed to quantify bioavailability, such as area under the curve (AUC), maximum drug concentration (C_{max}) and time at which maximum concentration is reached (T_{max}). Bioequivalence of the two formulations is typically assessed with respect to one or more of these summary measures rather than the profiles themselves. It is not difficult to contrive a situation in which two formulations have different profiles and identical summary measures, although some would argue that it would be extremely unusual for this to occur when AUC, C_{max} and T_{max} are all equal. For some drugs, it may be reasonable to assess bioequivalence with respect to one or another summary measures if the shape of the profile is not clinically important. However, for other drug classes it may be important that two formulations with different plasma concentration profiles not be considered bioequivalent even if they are similar with respect to some summary measures.

Several authors have proposed methods for comparing two formulations over the entire profile. Rescigno (1992) suggested a single index as a measure of profile

similarity:

$$\xi_m = \left(\frac{\sum_{j=1}^{n} w_j |y_{\mathrm{R}j} - y_{\mathrm{T}j}|^m}{\sum_{j=1}^{n} w_j |y_{\mathrm{R}j} + y_{\mathrm{T}j}|^m} \right)^{1/m} \tag{1}$$

where $y_{\mathrm{R}j}$ and $y_{\mathrm{T}j}$ are the observed drug concentrations at time t_j under the reference and test formulations respectively, w_j is a weight chosen to reflect the importance of the sampling time t_j, and m is a positive integer. The index takes its minimum value of zero when the profiles are identical at every sampling time, and its maximum value of one when one of the two profiles is zero at all sampling times. Choices for w_j and m could be based on pharmacologic and statistical considerations. However, from a practical point of view, a drawback of this approach is the lack of meaningful pharmacologic interpretation for ξ and the difficulty in determining which values of ξ should be deemed to represent bioequivalence.

Chinchilli and Elswick (1997) proposed a different method for summarizing profile differences. Their approach is somewhat complex, but can be thought of as an accumulation of the pointwise relative differences between the two profiles. Roughly speaking, their index measures the total relative difference, whereas Rescigno's index measures the relative total difference. A potential problem with this approach is that large relative differences at one or two time points can have undue influence on the overall measure, making it overly sensitive. From a practical point of view their measure suffers the same lack of pharmacologic interpretability as the Rescigno index although to a lesser extent.

An obviously naive approach would be a repeated measures analysis including a time by formulation interaction. However, this analysis does not lend itself to the usual bioequivalence testing framework which we briefly review in the next section. In addition, as with the Chinchilli and Elswick approach, it may be overly sensitive to clinically unimportant differences. In this paper, we propose a method for assessing bioequivalence over the entire profile which can be interpreted in familiar terms and represents a compromise between the insensitivity of summary measures and the oversensitivity of pointwise comparisons.

2. Area Under the Curve

The total area under the plasma drug concentration profile is an attractive summary measure for comparing drug bioavailabilities because it represents total exposure to the drug, and under certain conditions is proportional to the total amount of drug absorbed. For bioequivalence trials, in which plasma drug concentration is estimated at discrete times, the area under the curve is typically estimated by linear interpolation.

$$\mathrm{AUC} = \sum_{j=1}^{J} (t_j - t_{j-1}) \frac{y_j + y_{j-1}}{2} \tag{2}$$

where y_j is the observed plasma drug concentration at time t_j, $t_0 = y_0 = 0$, and t_J is the last sampling time. In cases where the plasma drug concentration has not returned to zero by the last sampling time (i.e., $y_J > 0$), it is common to estimate the unobserved portion of the curve under a model of first-order elimination kinetics:

$$\mathrm{AUC}_\infty = \mathrm{AUC} + y_j/K_e \tag{3}$$

where K_e is the first-order elimination rate estimated either by modeling the observed profile, or separate data collected from the same subject after intravenous drug administration. This is likely to be a reasonable adjustment if the drug has been completely absorbed by the last sampling time.

Bioequivalence typically is defined by an acceptance region for the relative difference between the AUC of the test (T) and reference (R) formulations.

$$\theta_l < \frac{\mathrm{AUC_T} - \mathrm{AUC_R}}{\mathrm{AUC_R}} < \theta_u \tag{4}$$

The United States Food and Drug Administration (FDA) currently requires $\theta_l = -0.2$ and $\theta_u = 0.2$ (which are symmetric on an additive scale) for analyses done on untransformed data. For log-transformed data the requirement is expressed as

$$\log(1+\theta_l) < \log(\mathrm{AUC_T}) - log(\mathrm{AUC_R}) < \log(1+\theta_u) \tag{5}$$

and the FDA uses $\theta_l = -0.2$ and $\theta_u = 0.25$ (which are symmetric on a multiplicative scale). The second pair of limits implies an acceptance region which is the same regardless of which formulation is viewed as the reference, however the first pair does not. Statistical inference in bioequivalence trials is typically cast as a hypothesis test of:

$$H_o:\ \frac{\mu_\mathrm{T} - \mu_\mathrm{R}}{\mu_\mathrm{R}} \le \theta_l \quad \text{or} \quad \frac{\mu_\mathrm{T} - \mu_\mathrm{R}}{\mu_\mathrm{R}} \ge \theta_u$$

versus

$$H_a:\ \theta_l < \frac{\mu_\mathrm{T} - \mu_\mathrm{R}}{\mu_\mathrm{R}} < \theta_u$$

where μ_R and μ_T are the population means of $\mathrm{AUC_R}$ and $\mathrm{AUC_T}$ respectively.

This type of test is commonly referred to as a test of *average* bioequivalence since the null and alternative hypotheses specify regions for the means only. If the null and alternative hypotheses are expanded to the entire distribution of the AUCs for each formulation, the resulting tests are called *population* bioequivalence tests. Berger and Hsu (1996) provide a thorough overview of these types of tests. Another approach attempts to demonstrate that equation 4 holds for each subject in the trial. This approach is termed *individual* bioequivalence, see Anderson and Hauck (1990), Sheiner (1992) and Schall (1995). AUC is the most commonly used measure of bioavailability, but bioequivalence testing of $\mathrm{C_{max}}$, $\mathrm{T_{max}}$ and other summary measures has been well characterized. In addition, multivariate tests of simultaneous bioequivalence of several summary measures have also been developed, see Berger and Hsu (1996).

AUC can be viewed as a family of bioavailability summary measures. We define

$$\mathrm{AUC}_k = \sum_{j=1}^{k} (t_j - t_{j-1}) \frac{y_j - y_{j-1}}{2} \tag{6}$$

so that $\mathrm{AUC}_J = \mathrm{AUC}$ when t_J is the last sampling time. For $k < J$, AUC_k is called partial AUC. Several authors have proposed the use of partial AUC in bioequivalence testing, see Martinez and Jackson (1991) and Niazi et al. (1997). This suggests the possibility of constructing an overall comparison of plasma concentration profiles in terms of partial AUCs. The left-hand panels in Figure 2 show the plasma concentration profiles from the three subjects highlighted in Figure 1. The middle panels are the corresponding partial AUC at each sampling time. The right-hand panels show the relative partial AUC at each sampling time, i.e., $(\mathrm{AUC}_{\mathrm{T}k} - \mathrm{AUC}_{\mathrm{R}k})/\mathrm{AUC}_{\mathrm{R}k}$. The dashed lines correspond to $\theta_l = -0.2$ and $\theta_u = 0.2$. The last point on the curve corresponds to the relative difference in total AUC which is commonly used to evaluate bioequivalence. The erratic behavior at the early sampling times reflects the fact that the reference profile is very low and $\mathrm{AUC}_{\mathrm{R}k} \approx 0$. It should also be noted that although in this example the relative partial AUC curves tend to flatten out at the end of the profile, this is due to the fact that both plasma concentration profiles have returned to zero. This might not occur if the plasma concentration of one or both formulations was non-zero at the last sampling time.

While this does provide a framework for comparing curves which can be viewed as a generalization of the usual AUC comparison, a major drawback is that it does not allow an independent comparison over different regions of the curves. For example, a difference over the last few hours of the profile could be masked if the formulations were similar up until that point. An absolute comparison of the observed plasma concentration profiles in a pointwise manner would remedy this problem, but as mentioned earlier does not lend itself to the usual definition of acceptance regions in terms of relative differences. A relative comparison of the profiles in a pointwise manner would behave erratically when the reference formulation profile is low. In addition, even though interpretation would be straightforward, a large relative difference in the first few minutes may not be clinically important.

One solution would be to make a pointwise relative comparison of the profiles over a predetermined window of clinical importance, e.g., when the reference profile is known to be greater than some threshold. Presumably, one would require that the relative difference between the profiles should fall within some acceptance region at each sampling time. In order to guarantee that this also would meet the usual bioequivalence requirement, it would be necessary to make the pointwise acceptance regions smaller than the usual FDA acceptance region. However, this requirement may be too stringent. Consider the situation in which the test formulation is lower than acceptance at the second sampling time, but is then higher than acceptance at the third sampling time. Since early sampling times typically are very close to each other this might not represent a clinically important difference, whereas the same sort of discrepancy at two later, more distant, sampling times might be clinically

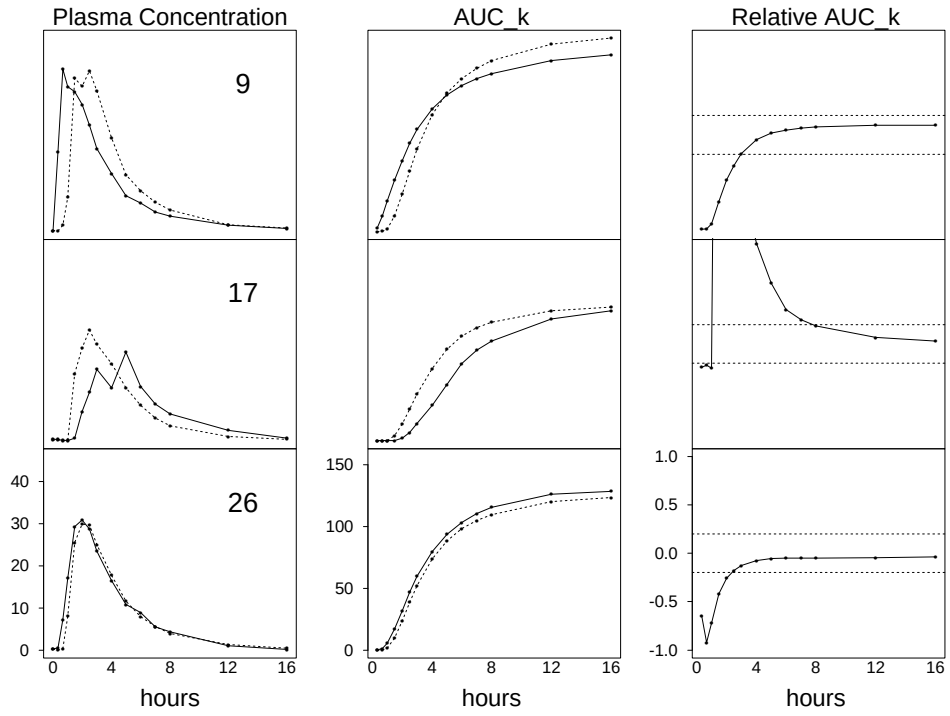

Figure 2: Plasma concentration profiles, partial AUC, and relative partial AUC; each row corresponds to one subject.

important. A general problem with pointwise comparisons is that they are design dependent.

A reasonable solution to this problem would be to assess formulation bioequivalence within predetermined regions of clinical importance. This can be viewed as a smoothing or averaging of the pointwise comparison, and is a compromise between the pointwise comparison which could be overly sensitive to clinically unimportant differences and the total AUC comparison which essentially averages over the entire profile and could be insensitive to clinically important differences. The approach we propose represents this type of compromise, where the regional smoothing is meant to be consistent with the usual comparison of relative AUC. In this derivation, we rewrite the usual AUC estimate in the following form:

$$\text{AUC} = \sum_{j=1}^{J} w_j y_j \tag{7}$$

where $w_j = (t_{j+1} - t_{j-1})/2$ for $j = 1, \ldots, J-1$ and $w_J = (t_J - t_{J-1})$. If w_J is replaced by $w_J^* = w_J + 1/K_e$, then the weighted sum is the usual estimate of

AUC_∞. The relative difference in AUC can be expressed as:

$$\frac{\mathrm{AUC_T} - \mathrm{AUC_R}}{\mathrm{AUC_R}} = \sum_{j=1}^{J} \frac{w_j(y_{\mathrm{T}j} - y_{\mathrm{R}j})}{\mathrm{AUC_R}} \tag{8}$$

with w_j defined as above. Substituting into the expression for the bioequivalence acceptance region (equation 4) gives:

$$\theta_l < \sum_{j=1}^{J} \frac{w_j(y_{\mathrm{T}j} - y_{\mathrm{R}j})}{\mathrm{AUC_R}} < \theta_u \tag{9}$$

This expression can be decomposed into mutually exclusive acceptance regions. One natural decomposition would be

$$\begin{aligned} \theta_l/J < \frac{w_1(y_{\mathrm{T}1} - y_{\mathrm{R}1})}{\mathrm{AUC_R}} < \theta_u/J \\ \theta_l/J < \frac{w_2(y_{\mathrm{T}2} - y_{\mathrm{R}2})}{\mathrm{AUC_R}} < \theta_u/J \\ \vdots \\ \theta_l/J < \frac{w_J(y_{\mathrm{T}J} - y_{\mathrm{R}J})}{\mathrm{AUC_R}} < \theta_u/J \end{aligned} \tag{10}$$

so that if all of these expressions hold, then expression 9 holds. The converse is not necessarily true, i.e., expression 9 does not imply that each component of expression 10 is true.

This provides a pointwise assessment of the relative difference of the profiles with respect to total AUC. That is, it represents the contribution of each sampling time to the relative difference in total AUC. An advantage of this over a pointwise assessment of the relative difference of the profiles (i.e., $(y_{\mathrm{T}j} - y_{\mathrm{R}j})/y_{\mathrm{T}j}$) is that it is not ill-behaved when $y_{\mathrm{T}j}$ is small at the ends of the profile. The pharmacological interpretation of this kind of comparison is that relative differences at specific points in the profile should be assessed in light of the total exposure to the drug. The middle set of panels in Figure 3 show the contribution to relative AUC for each sampling time for the three profiles in our example. The dashed lines represent the pointwise acceptance region corresponding to $\theta_l/J = -0.2/14$ and $\theta_u/J = 0.2/14$. For comparison, the right-hand panels are the same as in Figure 2. The first profile (subject number 9) provides an example of a situation in which the curves are different over the middle of the profile, but the relative partial AUC_k does not reflect this due to a difference in the opposite direction over the early part of the profile. As mentioned earlier however, this approach may be overly sensitive to clinically unimportant differences.

A much more gross assessment of the relative difference can be achieved by

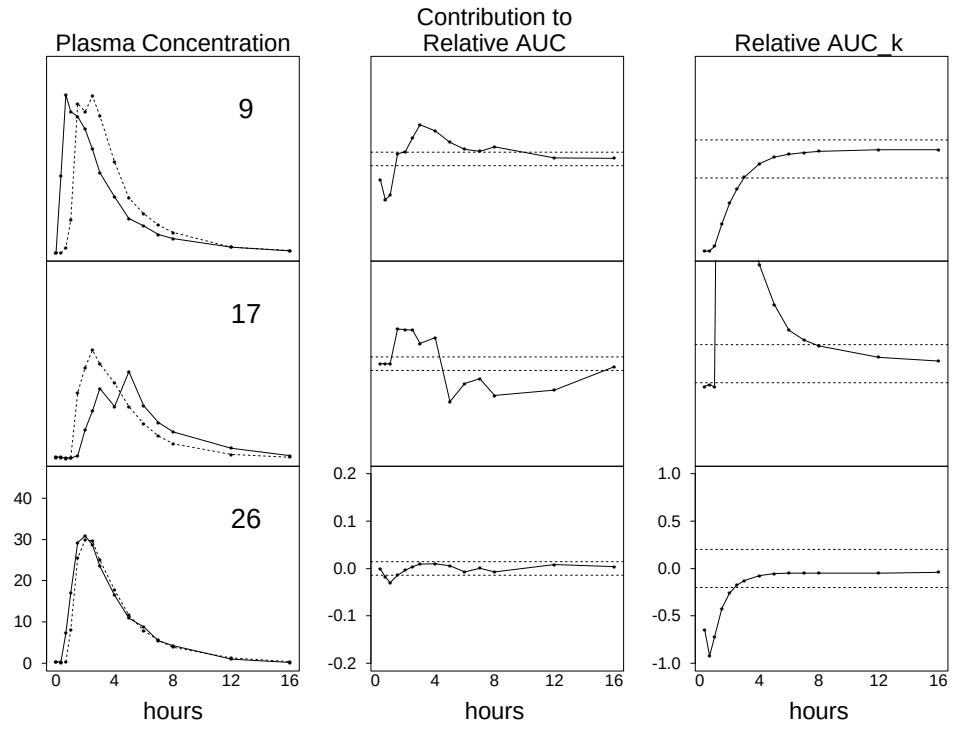

Figure 3: Plasma concentration profiles, pointwise contribution to relative AUC, and relative partial AUC; each row corresponds to one subject.

decomposing the time course into halves.

$$\begin{aligned} \theta_l/2 &< \sum_{j:t_j<t_J/2} \frac{w_j(y_{\mathrm{T}j} - y_{\mathrm{R}j})}{\mathrm{AUC_R}} < \theta_u/2 \\ \theta_l/2 &< \sum_{j:t_j>t_J/2} \frac{w_j(y_{\mathrm{T}j} - y_{\mathrm{R}j})}{\mathrm{AUC_R}} < \theta_u/2 \end{aligned} \tag{11}$$

Alternatively, the time course could be divided into quarters. The middle set of panels in Figure 4 shows the contribution to the relative AUC for each half of the time course. The dashed lines represent an acceptance region defined by $\theta_l/2 = -0.2/2$ and $\theta_u/2 = 0.2/2$. In this case subject number 9, who appeared to have fairly different profiles, is shown to be bioequivalent over both the first and the second half of the time course. The right-hand set of panels shows the relative difference between the profiles when dividing the time course into quarters with dashed lines representing an acceptance region defined by $\theta_l/4 = -0.2/4$ and $\theta_u/4 = 0.2/4$. Here we see that subject number 17, for whom the reference profile appeared to be too high over the first eight hours (middle panel), exhibits an interesting pattern in which the reference profile appears to be too high during the first four hours and

too low during the second four hours.

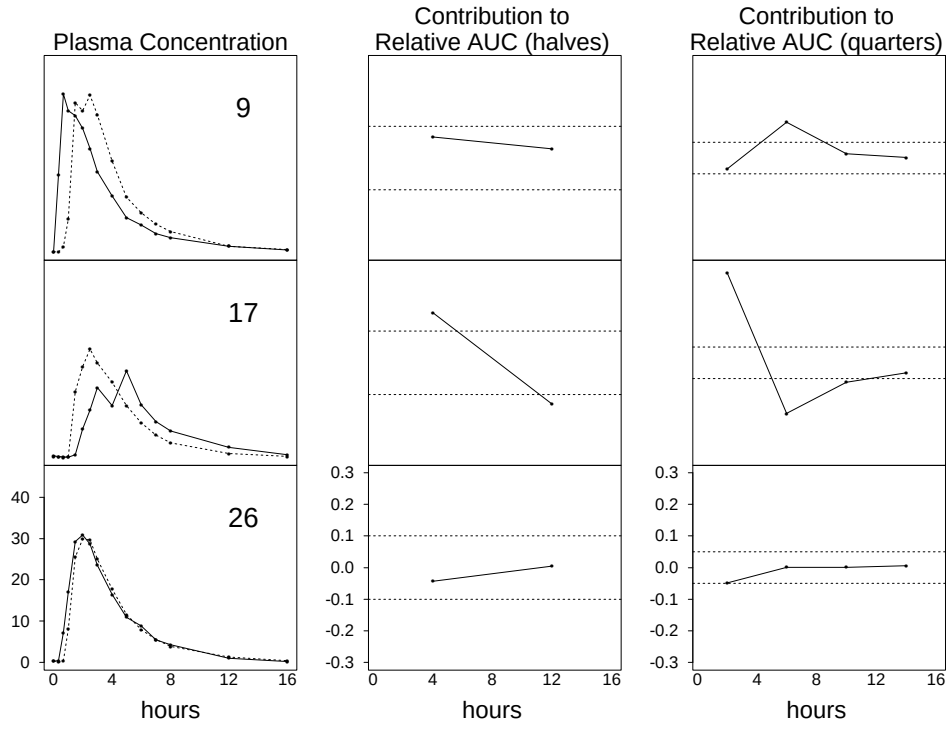

Figure 4: Plasma concentration profiles, regional contribution to relative AUC (8 hour window), regional contribution to relative AUC (4 hour window); each row corresponds to one subject.

In this example we have divided the time course into halves and quarters arbitrarily for illustrative purposes. If one is concerned that the usual relative AUC assessment of bioequivalence could miss a clinically important difference for a particular drug, then presumably it would be reasonable to predefine regions of the time course in which clinically important differences could appear. In this example we also have arbitrarily divided the acceptance region into equal parts. In practice it would be reasonable to allow different acceptance regions for different portions of the time course.

3. **Example**

The following table summarizes the results of an analysis of all 26 subjects in the ibuprofen study. The first four columns show the contribution of each four-hour block to the relative difference in total AUC. The fifth and sixth columns show the contribution of each eight-hour block to the relative difference in total AUC. The

last two columns show the relative difference in total AUC and C_{max}. Within each row the sum of the first four columns equals the relative difference in total AUC. The same is true for the sum of the fifth and sixth columns. Boldface entries denote values that fall outside the bioequivalence acceptance region defines as $\pm 0.2/4$ for the first four columns, $\pm 0.2/2$ for the fifth and sixth columns, and ± 0.2 for the last two columns.

Comparison of summary statistics for evaluating bioequivalence.

	Contribution to Relative AUC							
	4 hour blocks				8 hour blocks		Relative	Relative
Subj	0-4	4-8	8-12	12-16	0-8	8-16	AUC	C_{max}
1	0.039	0.013	-0.022	-0.016	0.062	-0.050	0.012	-0.108
2	0.019	**-0.070**	-0.004	0.004	-0.044	-0.006	-0.050	-0.115
3	0.030	-0.025	-0.002	0.001	0.008	-0.004	0.004	0.063
4	0.040	**0.064**	0.012	0.010	**0.104**	0.021	0.124	0.146
5	-0.039	**0.150**	0.013	-0.006	0.096	0.018	0.114	-0.138
6	-0.043	0.007	0.004	0.007	-0.035	0.012	-0.022	-0.100
7	**-0.114**	**0.175**	0.024	0.005	0.045	0.042	0.087	-0.179
8	0.020	**0.130**	**0.051**	0.017	**0.117**	0.099	**0.216**	-0.079
9	-0.034	**0.114**	0.014	0.002	0.067	0.029	0.096	-0.008
10	**-0.074**	**0.080**	0.000	-0.009	-0.004	0.002	-0.002	**-0.238**
11	0.016	0.002	-0.001	-0.002	0.018	-0.002	0.016	0.194
12	**0.100**	0.027	-0.001	-0.002	**0.124**	0.000	0.125	**0.219**
13	0.038	0.038	0.007	0.004	0.070	0.021	0.091	0.034
14	-0.048	0.010	0.003	0.004	-0.035	0.001	-0.034	-0.182
15	**0.065**	-0.004	-0.002	-0.006	0.064	-0.019	0.045	**0.365**
16	-0.015	-0.020	0.004	0.006	-0.032	0.007	-0.025	-0.039
17	**0.285**	**-0.160**	**-0.062**	-0.031	**0.159**	**-0.130**	0.029	**0.248**
18	**-0.139**	0.044	-0.001	-0.001	-0.096	-0.001	-0.097	**-0.457**
19	0.019	**-0.089**	-0.020	-0.005	-0.058	-0.033	-0.091	0.187
20	-0.036	0.024	0.001	-0.001	-0.014	0.004	-0.010	**-0.297**
21	**0.297**	**-0.100**	0.015	0.015	**0.193**	0.039	**0.231**	0.177
22	0.009	0.017	0.001	0.001	0.027	0.001	0.028	0.049
23	**-0.188**	**0.159**	0.024	0.012	-0.048	0.060	0.013	0.108
24	**-0.161**	**0.069**	0.009	0.004	-0.097	0.018	-0.079	**-0.242**
25	0.024	0.011	-0.004	-0.005	0.036	-0.012	0.023	-0.156
26	-0.048	0.001	0.001	0.006	-0.044	0.005	-0.039	-0.029

Boldface entries indicate values falling outside the acceptance region.

Only 2 of the 26 subjects fell outside the bioequivalence acceptance region with respect to total AUC, another 7 failed with respect to C_{max}. Looking at the first four columns, 15 subjects failed with respect to at least one of the four-hour blocks, 7 of these correspond to subjects who did not fail on either AUC or C_{max} (subjects number 2, 4, 5, 7, 9, 19 and 23). Based on Figure 1, the most notable of these 6 is subject number 23 who exhibited a marked difference in absorption rate for the two formulations. Looking at the fifth and sixth columns, 5 subjects failed with respect to at least one of the eight-hour blocks. Three of these, subjects number 4, 12 and 17, did not fail with respect to total AUC. Only two subjects exhibited any formulation difference over the last eight hours of the study, and for most subjects the plasma drug concentration was very low during this time period.

From our perspective, four of the most interesting rows in the table correspond to subjects number 7, 10, 23 and 24. These subjects did not fail with respect to

total AUC or either of the eight-hour blocks, but did fail with respect to both of the first two four-hour blocks. They are interesting because they failed (either too high or too low) in the first block, but then failed in the opposite direction in the second block. Such a pattern cannot be found by looking at more than one partial AUC because the differences cancel. This leaves us with the important question, which cannot be answered on statistical grounds, of whether or not such a difference is clinically important.

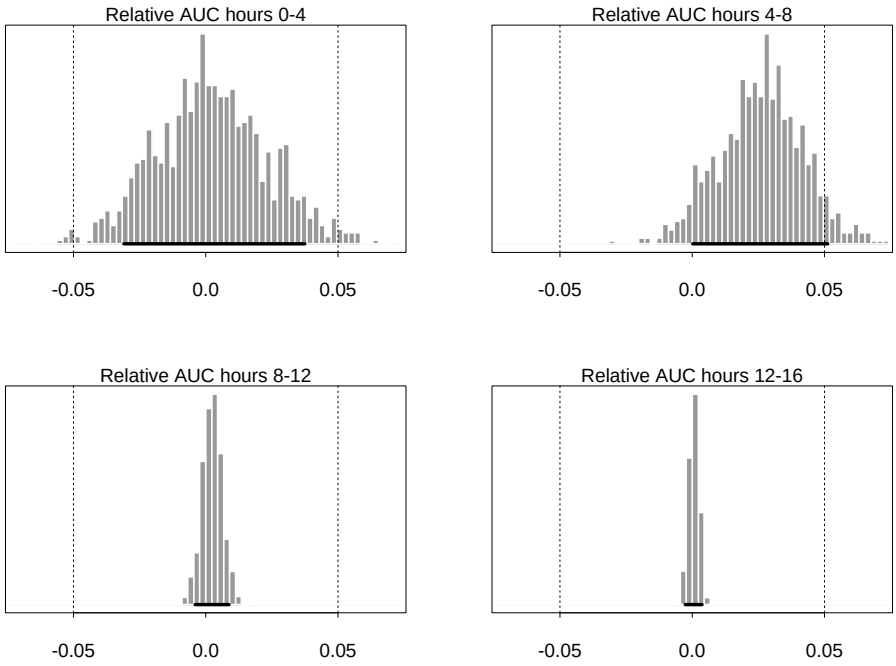

Figure 5: Bootstrap distributions of mean regional contribution to relative AUC; horizontal lines denote bioequivalence acceptance region, heavy vertical lines denote middle 90% of bootstrap distribution.

In addition, we generated bootstrap distributions for the mean contribution of each four-hour block to the relative difference in total AUC, illustrated in Figure 5. The bottom two panels show that the mean contribution to the relative AUC over the last two four-hour blocks lies well within the acceptance region defined by $(-0.2/4, 0.2/4)$. The middle 90% of the bootstrap distribution for the mean contribution of the first four-hour block (top left panel) lies entirely within the acceptance region, but it appears that there is a large amount of variation associated with the relative AUC during this time period. The middle 90% of the bootstrap distribution for the mean contribution of the second four-hour block (top right panel) does not lie entirely within the acceptance region. Therefore, even though the

mean of the distribution does lie within the acceptance region, we would conclude that the two formulations are not bioequivalent during this time period. It should be noted that if we had divided up the total acceptance region such that more tolerance was given to the early part of the curve, this analysis would have concluded that the two formulations are bioequivalent.

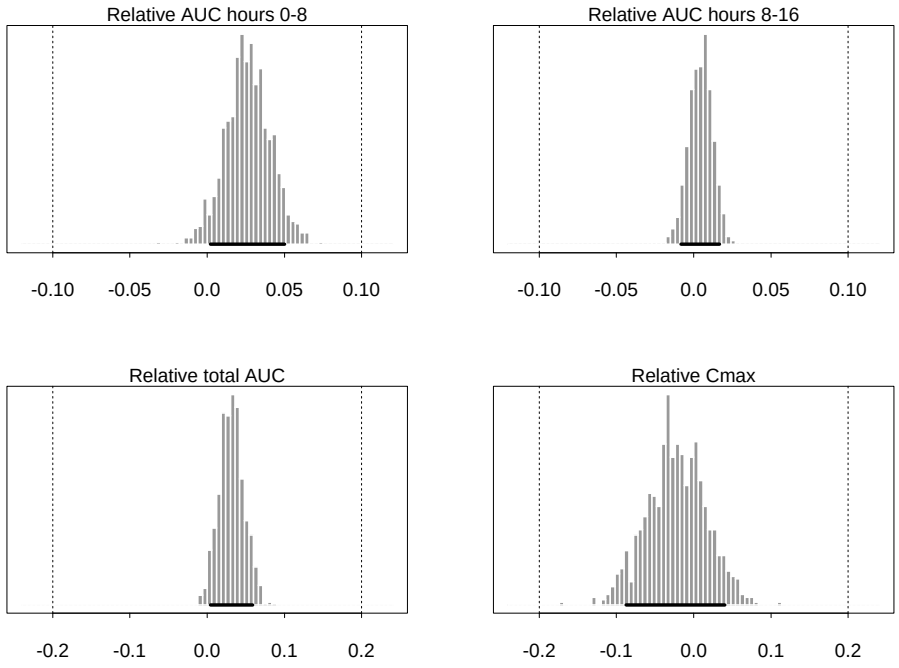

Figure 6: Bootstrap distributions of mean regional contribution to relative AUC, relative total AUC, and relative C_{max}; horizontal lines denote bioequivalence acceptance region, heavy vertical lines denote middle 90% of bootstrap distribution.

We also generated bootstrap distributions for the mean contribution of each eight-hour block to the relative difference in total AUC, illustrated in the top panels of Figure 6. Even though the mean contribution of the first eight-hour block (top left panel) is larger than zero, the bootstrap distribution lies entirely within the acceptance region defined by $(-0.2/2, 0.2/2)$. As expected, the distribution of the mean contribution of the second eight-hour block (top right panel) lies well within the acceptance region. The bootstrap distributions for the mean relative difference in total AUC and C_{max} (bottom panels) also lie entirely within the usual acceptance region.

4. Discussion

Although we only addressed average bioequivalence in our example, the bootstrap approach easily can be adapted to obtain confidence intervals for various percentiles of the relative regional AUCs. It also would be straightforward to develop test statistics for assessing bioequivalence with respect to relative regional AUCs. Total AUC is commonly analyzed after logarithmic transformation; however, that is not possible for these measures. Perhaps other formulations of regional profile differences based on this idea would be more suitable for normal theory-based test statistics. Another issue to be addressed is the optimal strategy for decomposing the acceptance region.

In summary, the approach to assessing bioequivalence we have proposed can be interpreted in terms familiar to the pharmacologist and consistent with current FDA requirements. Since it is subject specific, it can be used with average, population, and individual bioequivalence inference methods. One very important feature of this approach is that the clinically relevant regions of the time course must be specified a priori if it is to be used in a regulatory setting. In a purely exploratory analysis, the utility of this approach is that it can provide a smoothed picture of the pointwise relative difference between the profiles.

References

Anderson S, Hauck WW. (1990). Consideration of individual bioequivalence. *Jour. Pharmacokinetics and Biopharmaceutics*, 18:259-273.

Berger RL, Hsu JC. (1996). Bioequivalene trials, intersection-union tests and equivalence confidence sets. *Statistical Science*, 11(4):283-319.

Chinchilli VM, Elswick RK. (1997). The multivariate assessment of bioequivalence. *Jour. Biopharmaceutical Statistics*, 7(1):113-123.

Martinez MN, Jackon AJ. (1991). Suitability of various noninfinity area under the plasma concentration-time curve (AUC) estimates for use in bioequivalence determinations: relationship to AUC from zero to time inifinity (AUC0-INF). *Pharmaceutical Research*, 8(4):512-517.

Niazi SK, Alam SM, Ahmad SI (1997). Partial-area method in bioequivalence assessment: naproxen. *Biopharmaceutics and Drug Disposition*, 18(2):103-116.

Rescigno, A (1992). Bioequivalence. *Pharmaceutical Research*, 9(7):925-928.

Schall R (1995). Assessment of individual and population bioequivalence using the probability that bioavailabilities are similar. *Biometrics*, 51:615-626.

Sheiner LB (1992). Bioequivalence revisited. *Statistics in Medicine*, 11:1777-1788.

David T. Mauger and Vernon M. Chinchilli
The Pennsylvania State University
Department of Health Evaluation Sciences, H173
500 University Drive
Hershey, PA 17033
U.S.A.
E-mails: dmauger@psu.edu, vchinchi@psu.edu