Novel Direct Curve Comparison Metrics for Bioequivalence

James E. Polli[1,3] and Angus M. McLean[2]

Received January 3, 2001; accepted March 7, 2001

Purpose. The object of this work was to devise four new direct curve comparison (DCC) metrics and examine each metric's distribution properties and performance characteristics.

Methods. DCC metrics, Cmax, and AUCi were calculated from two bioequivalence studies of three sustained release carbamazepine formulations, where a range of profile similarity was observed. DCC metric values and their confidence intervals were compared to Cmax and AUCi.

Results. The DCC metrics ρ, ρ_m, δ_a, and δ_s exhibited more favorable distributions than Cmax and AUCi ratios, which were frequently skewed. The DCC metrics performed differently than Cmax and AUCi ratios in profile comparisons due to the nature of the DCC metrics. Unlike Cmax and AUCi, the DCC metrics utilize all data points to directly compare entire profiles. Each DCC metric appears to measure "exposure" in a single assessment. Possible bioequivalence acceptance criteria are: $\rho \leq 1.40$, $\rho_m \leq 0.35$, $\delta_a \leq 0.27$, and $\delta_s \leq 0.102$.

Conclusions. These DCC metrics, particularly ρ_m, are promising bioequivalence metrics for "exposure."

KEY WORDS: bioequivalence; bioavailability; carbamazepine.

INTRODUCTION

The review of bioequivalence data involves a graphic comparison of the test and reference cross-over plasma concentration-time profiles. In performing this review, qualitative impressions are formed of the differences in maximum plasma concentration, the extent of absorption, and the shape of plasma concentration-time profiles. Presently, the quantification of bioequivalence is based on differences in the maximum plasma concentration (Cmax) and differences in the area under the plasma concentration-time profile extrapolated to infinity (AUCi). Neither of these metrics compares the entire shapes of the two profiles, although a metric that compares overall profile shape would arguably be valuable. Such a metric can be described as a direct curve comparison metric (DCC), because it, unlike Cmax and AUCi, would directly compare entire test and reference profiles. Rescigno (1) and Chinchilli and Elswick (2) have suggested such bioequivalence metrics. Unlike Cmax and AUCi, a DCC approach utilizes all data points, compares profiles at the same time points, and provides a single evaluation. Relative to Cmax and AUCi, DCC metrics better detect curve shifts (i.e. drug absorption lag times) and differences between multiple peak profiles, as is often the case for extended release products (3).

The object of this work was to devise four new DCC metrics and examine each metric's distribution properties and performance characteristics. This objective was undertaken from the viewpoint that Cmax and AUCi do not sufficiently compare the shapes of plasma concentration-time profiles in all cases. The goal in developing these DCC metrics was to explore the potential for such a metric to supplement, or indeed replace, Cmax and AUCi as bioequivalence metrics.

Limitations of Cmax and AUCi as bioequivalence metrics have previously been identified, and in part arise from the regulatory and legal definitions concerning bioequivalence and bioavailability. Bioequivalent drug products must display comparative bioavailability when studied under similar experimental conditions (4). Bioavailability is defined as "the rate and extent to which the active drug ingredient or therapeutic ingredient is absorbed from a drug product and becomes available at the site of drug action" (5). This definition, which emphasizes rate and extent of drug absorption, has resulted in the use of the maximum plasma concentration (Cmax) and the area under the plasma concentration-time profile extrapolated to infinity (AUCi) as bioequivalence metrics. Cmax and AUCi serve as metrics for rate and extent, respectively.

While AUCi exhibits favorable properties as a metric for extent (6), Cmax has been criticized as a metric for rate. Cmax reflects extent (6–8). These properties result in the evaluation of extent twice, once from AUCi and once from Cmax, with Cmax the generally more variable metric. While not an ideal metric for rate, Cmax can be valuable as a metric for dose-dumping, a safety concern (7). In Canada, for uncomplicated drugs, the Cmax ratio needs lie between 80–120%, and not its confidence interval (9).

For average bioequivalence, one approach to address the limitations of Cmax is the use of an alternative metric for rate. Cmax/AUC (10–14) and partial AUC (11,14,15) have been examined as rate metrics, but opposing recommendations have emerged. A method for Tmax (16) and other novel rate metrics (17) have also been suggested. Cmax/AUC has also been suggested, in the context as a secondary metric to AUC (18,19), but opposing recommendations have emerged (20,21).

Another approach to address the limitations of Cmax is the revisitation of the definitions of bioavailability and bioequivalence. Tozer *et al.* (22) encourage the concept of "exposure" rather than "rate and extent of absorption," and indicate the goal of bioequivalence testing to be the assurance of similar concentration-time profiles. Since infinite profiles can calculate to the same AUCi, AUCi is not an acceptable single exposure metric (11,22).

In an effort to identify potential "exposure" metrics (i.e. metrics that compare entire plasma profiles), the object of this work was to devise four new DCC metrics and examine each metric's distribution properties and performance characteristics. The DCC metrics were applied to plasma concentration-time profiles data sets from two bioequivalence studies of three sustained release carbamazepine formulations. The data sets were carefully selected since marginal bioequivalence and marginal bioinequivalence, as dictated by Cmax

[1] School of Pharmacy, University of Maryland, Baltimore, Maryland 21201.

[2] Shire Laboratories, Rockville, Maryland 20850.

[3] To whom correspondence should be addressed at Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 North Pine Street, Baltimore, Maryland 21201. (e-mail: jpolli@rx.umaryland.edu)

0724-8741/01/0600-0734$19.50/0

and AUCi, was evident in the data sets. DCC confidence intervals and possible bioequivalence criteria are provided.

THEORETICAL

Four new metrics for bioequivalence (ρ, ρ_m, δ_a, and δ_s) are defined below. The basis for each metric's construction is provided. In an effort to compare entire profiles, each metric utilizes all plasma profile data and reflects entire profile similarity/dissimilarity. In the final form of each metric, the sum of the test and reference concentrations serves as a weight to give importance to the higher concentrations. Each metric is constructed such that the assignment of the test and reference product is without consequence.

Rho

ρ considers the ratio of the profiles at the same time points. To allow all points of the profiles to be included in a ratio approach, ρ utilizes the larger of either T/R or R/T, where R and T are the concentrations of the reference and test products, respectively, at time t. For the unweighted comparison of a single pair of plasma concentrations at time t, $\rho^u = RATIO$ where $RATIO$ is the larger of T/R and R/T. The superscript u denotes that the metric is unweighted. The larger, rather than smaller, ratio is used since the resulting distribution is likely to be skewed positively (to the right), which may be converted to a normal distribution via ln transformation. As discussed in the Methods section, the ln transformation approach in current regulatory use for Cmax and AUCi analysis is followed here.

For n number of paired plasma samples, the unweighted metric is $\rho^u = 1/n\ \Sigma_{t=1}^{n}\ RATIO_t$, where $RATIO$ is the larger of T/R and R/T. The general expression for a weighted arithmetic mean is $\overline{X} = \Sigma_{t=1}^{n} w_t X_t / \Sigma_{t=1}^{n} w_t$. Applying a weight of $(R + T)$ to $\rho^u = RATIO$ gives the weighted metric:

$$\rho = \frac{\sum_{t=1}^{n}(R_t + T_t) \times RATIO_t}{\sum_{t=1}^{n}(R_t + T_t)} \tag{1a}$$

It should be noted that $\rho \geq 1$. If either R or T (but not both) are equal to zero, then a value of 10 is assigned to $RATIO$, although other solutions could be applied. If both R and T are equal to zero, then a value of 1 is assigned to $RATIO$.

Since a natural logarithm approach was used to construct a confidence interval for ρ (see Methods section), the numerator and denominator of ρ are denoted as N_ρ and D, respectively.

$$N_\rho = \sum_{t=1}^{n}(R_t + T_t) \times RATIO_t \tag{1b}$$

$$D = \sum_{t=1}^{n}(R_t + T_t) \tag{2}$$

Because the dissimilarity of points at low concentrations contributes to overall curve dissimilarity, an entire profile comparison approach will be influenced by the analytical limit of quantification (LOQ). For example, the value of ρ may increase or decrease when test and reference points are made quantifiable through a more assay. While this attribute may be undesirable, its impact is attenuated by the $(R + T)$ weight.

Rho$_m$

The metric ρ_m also considers the ratio of plasma profiles. While ρ in Eq. 1a considers $RATIO$, ρ_m considers $RATIO - 1$. When weighted by $(R + T)$,

$$\rho_m = \frac{\sum_{t=1}^{n}(R_t + T_t) \times [RATIO_t - 1]}{\sum_{t=1}^{n}(R_t + T_t)} \tag{3a}$$

Hence, $\rho_m \geq 0$. For non-identical profiles that are less than 2-fold dissimilar, $\rho_m < 1 < \rho$.

The numerator of ρ_m is denoted as N_{ρ_m} and is

$$N_{\rho_m} = \sum_{t=1}^{n}(R_t + T_t) \times [RATIO_t - 1] \tag{3b}$$

The denominator is the same as Eq. 2.

Delta$_a$

δ_a (and δ_s below) considers the difference between two profiles relative to the size of the profiles. Using the absolute value of the difference for the numerator and the mean for the denominator yields $\delta_a^u = |R - T|/0.5(R + T)$, for the unweighted comparison of a single pair of plasma concentrations at time t. The subscript a denotes that the absolute difference is employed. For n number of paired plasma samples, $\delta_a^u = 1/n[|R_1 - T_1|/0.5(R_1 + T_1) + |R_2 - T_2|/0.5(R_2 + T_2) + \cdots |R_n - T_n|/0.5(R_n + T_n)] = 2/n\ \Sigma_{t=1}^{n} |R_t - T_t|/(R_t + T_t)$. Applying a weight of $(R + T)$ to δ_a^u gives the weighted metric

$$\delta_a = \frac{2\sum_{t=1}^{n}|R_t - T_t|}{\sum_{t=1}^{n}(R_t + T_t)} \tag{4a}$$

It should be noted that $0 \leq \delta_a$. Previously, Rescigno (1) defined the metric $\xi_1 = \Sigma_{t=1}^{n} |R_t - T_t|/\Sigma_{t=1}^{n}(R_t + T_t)$. δ_a is simply two-fold ξ_1.

The numerator of δ_a is denoted as N_{δ_a} and is

$$N_{\delta_a} = 2\sum_{t=1}^{n}|R_t - T_t| \tag{4b}$$

The denominator is the same as Eq. 2.

Delta$_s$

δ_s is similar in development to δ_a. The difference between two profiles relative to the size of the profiles is considered. Since squaring is a common method to avoid the absolute value function, the squared difference is used in the numerator, and the mean profile is used in the denominator, yielding $\delta_s^u = (R - T)^2/[0.5(R + T)]^2$ for the unweighted comparison of a single pair of plasma concentrations at time t. The subscript s denotes that the squared difference is employed.

For n number of paired plasma samples, $\delta_s^u = 1/n[(R_1 - T_1)^2/[0.5(R_1 + T_1)]^2 + (R_2 - T_2)^2/[0.5(R_2 + T_2)]^2 + \cdots (R_n - T_n)^2/[0.5(R_n + T_n)]^2] = 4/n \Sigma_{t=1}^{n} (R_t - T_t)^2/(R_t + T_t)^2$. Applying a weight of $(R + T)$ to $\delta_s^u = (R - T)^2/[0.5(R + T)]^2$ gives the weighted metric:

$$\delta_s = \frac{4\sum_{t=1}^{n} \frac{(R_t - T_t)^2}{(R_t + T_t)}}{\sum_{t=1}^{n}(R_t + T_t)} \tag{5a}$$

The numerator of δ_s is denoted as N_{δ_s} and is

$$N_{\delta_s} = 4\sum_{t=1}^{n} \frac{(R_t - T_t)^2}{(R_t + T_t)} \tag{5b}$$

The denominator is the same as Eq. 2.

METHODS

Carbamazepine Data Sets

Carbamazepine plasma profile data from two bioequivalence studies of three different sustained release formulations were used to study the distribution and performance characteristics of the four novel bioequivalence metrics (23). These studies were selected since results showed various profiles to be bioequivalent and bioinequivalent to one another, including some sets to be marginally bioequivalent. In study 1 formulations FAST and MODERATE-A were administered to 12 fasted, healthy volunteers in a two-way crossover study. Plasma profiles of formulations FAST and MODERATE-A are plotted in Fig. 1a.

In study 2 formulations of FAST, MODERATE-A, and MODERATE-B were administered to 12 fasted, healthy volunteers in a three-way crossover study. FAST and MODERATE-A were the same formulations in study 1. Plasma profiles are plotted in Fig. 1b.

These formulations were selected to characterize the new metrics since various formulation pairings exhibit differing degrees of profile similarity. Cmax from the formulations was FAST >> MODERATE-A ≥ MODERATE-B. Table I lists the conventional Cmax and AUCi bioequivalency results, from the ln transformation approach (24). Formulations MODERATE-A and MODERATE-B were similar. FAST differed modestly from MODERATE-A; FAST and MODERATE-A were bioequivalent in study 1, but bioinequivalent in study 2. FAST and MODERATE-B were bioinequivalent. A difference between MODERATE-A and MODERATE-B, which were similar in formulation design, was one subject who showed low Cmax and AUCi from a dose of MODERATE-B (about two-fold less).

Bioequivalence Metrics and Their Confidence Limits

Four new metrics (ρ, ρ_m, δ_a, and δ_s) were evaluated. They are defined above in Eq. 1a, 3a, 4a, and 5a, respectively. Cmax and AUCi ratios were also evaluated. The 90% confidence intervals for Cmax and AUCi ratios were determined (24). Using the subject-within-sequence mean-square error as the variance, ANOVA analysis indicated no sequence effect on lnCmax or lnAUCi ($p > 0.5$).

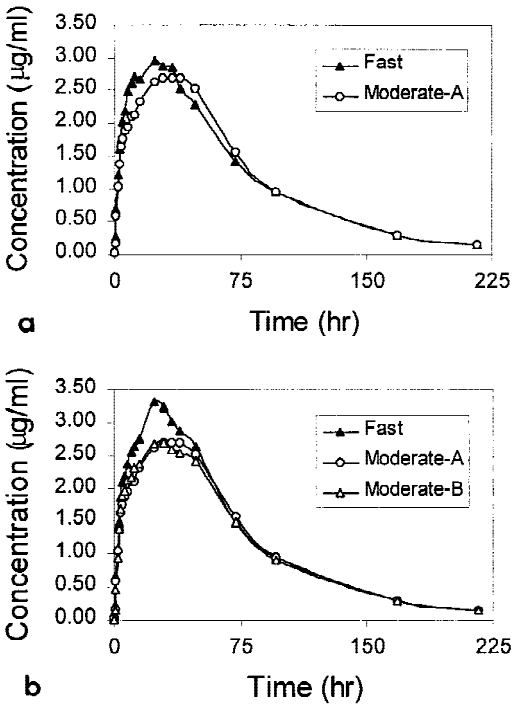

Fig. 1. Mean carbamazepine plasma profiles from study 1 and study 2. (a) In study 1, FAST and MODERATE-A profiles were modestly dissimilar (and bioequivalent). (b) In study 2, MODERATE-A and MODERATE-B profiles were similar, FAST and MODERATE-A profiles were modestly dissimilar (and bioinequivalent); and FAST and MODERATE-B were markedly dissimilar.

The upper 95% confidence limit for ρ, ρ_m, δ_a, and δ_s were determined using the same approach applied to Cmax and AUCi ratios. Briefly, like Cmax and AUCi ratios, this approach relied upon the identity $\ln N - \ln D = \ln (N/D) = \ln(metric)$. For each individual, the difference between the ln transformed numerator [ln$N$] and ln transformed denominator [ln$D$] was computed, and interpreted to be ln($metric$). For example, for ρ_m, $\ln N_{\rho_m} - \ln D$ was taken to yield $\ln\rho_m$. From the individual subjects, $\ln(\overline{metric})$ was calculated. The upper 95% confidence limit of $\ln(\overline{metric})$ was computed, using $s^2_{\ln(\overline{metric})}$ as the intrasubject variance:

$$s^2_{\ln(\overline{metric})} = \frac{MSE_{\ln N}}{n_{\ln N}} + \frac{MSE_{\ln D}}{n_{\ln D}} \tag{6}$$

where $MSE_{\ln N}$ and $MSE_{\ln D}$ are ANOVA mean-square error of ln$N$ and ln$D$, respectively; and $n_{\ln N}$ and $n_{\ln D}$ are sample size for ln$N$ and ln$D$, respectively.

As is done with Cmax and AUCi, Eq. 6 indicates that the variance of $\overline{\ln N} - \overline{\ln D}$ is the sum of the variance of $\overline{\ln N}$ and the variance of $\overline{\ln D}$, which assumes $\overline{\ln N}$ and $\overline{\ln D}$ are independent. Since $\overline{\ln N}$ and $\overline{\ln D}$ utilize the same data, $\overline{\ln N}$ and $\overline{\ln D}$ may exhibit interdependence. However, this was not the case (data not shown).

Like for lnCmax and lnAUCi, no sequence effect was evident for any of the novel bioequivalence metrics (ANOVA $p > 0.15$).

Additionally, the metric f_2, which is used to compare dissolution profiles, was applied to the carbamazepine data sets (25).

Table I. Novel Bioequivalence Metrics: Numerical Values and Comparison to Conventional Cmax and AUCi Pass/Fail Results[a,b]

Formulations	Cmax ratio	AUCi ratio	ρ	ρ_m	δ_a	δ_s
MODERATE-A vs.	1.02	1.02	1.26	0.217	0.178	0.0500
MODERATE-B	(0.91– 1.15)	(0.92– 1.12)	(≤1.34)	(≤0.288)	(≤0.225)	(≤0.0783)
(study 2)	Pass	Pass	Pass	Pass	Pass	Pass
FAST vs.	1.13	1.04	1.30	0.289	0.218	0.0751
MODERATE-A	(1.07– 1.20)	(0.99– 1.10)	(≤1.44)	(≤0.362)	(≤0.263)	(≤0.103)
(study 1)	Pass	Pass	Fail	Fail	Pass	Fail
FAST vs.	1.17	1.07	1.29	0.261	0.214	0.0647
MODERATE-A	(1.05– 1.32)	(0.97– 1.18)	(≤1.37)	(≤0.346)	(≤0.271)	(≤0.101)
(study 2)	Fail	Pass	Pass	Pass	Fail	Pass
FAST vs.	1.20	1.09	1.36	0.301	0.225	0.0794
MODERATE-B	(1.07– 1.35)	(0.99– 1.20)	(≤1.45)	(≤0.399)	(≤0.285)	(≤0.124)
(study 2)	Fail	Pass	Fail	Fail	Fail	Fail

[a] Least squares mean and 90% confidence interval (Cmax ratio and AUCi ratio) or upper 95% confidence limit (ρ, ρ_m, δ_a, and δ_s). Formulation comparisons are listed from most similar to least similar, according to Cmax.

[b] Passing or failing for ρ, ρ_m, and δ_a, and δ_s were based the following acceptance criteria: $\rho \le 1.40$, $\rho_m \le 0.35$, $\delta_a \le 0.27$, and $\delta_s \le 0.102$. Using the upper confidence limit, there was concordance between δ_a and the traditional criteria (Cmax and AUCi) in passing or failing studies. Meanwhile, ρ, ρ_m, and δ_s provided opposite results from the traditional criteria for the FAST vs MODERATE-A studies.

$$f_2 = 50 \log\left\{\left[1 + \frac{1}{n}\sum_{t=1}^{n}(R_t - T_t)^2\right]^{-0.5} \times 100\right\}$$

Skewness of Bioequivalence Metrics

In assessing the distribution properties of the new bioequivalence metrics, the distribution of each ρ, ρ_m, δ_a, and δ_s were examined by calculating the skewness. Skewness was determined from (26):

$$g1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}}$$

where $g1$ is skewness, n is the number of crossed-over plasma profiles, x_i is the metric value of sample i, and $\bar{x}_i$ is the mean metric value.

A data set is skewed right if data stretch to the positive direction. A data set is skewed left if data stretch to the negative direction. Positive values of $g1$ indicate positive skewness. Negative values of $g1$ indicate negative skewness. For perfectly symmetrical distribution, $g1 = 0$. A parametric confidence interval approach, such as the regulatory approach for Cmax and AUCi ratios (24) and the approach here for ρ, ρ_m, δ_a, and δ_s, assumes a normal distribution. Thus, a desirable tendency of a bioequivalence metric is $g1 = 0$. Skewness is a recommended method to assess whether a data set exhibits approximate normal distribution (26,27).

RESULTS AND DISCUSSION

Cmax and AUCi Ratios

Table I lists least squares mean Cmax and AUCi ratios for the four formulation comparisons. In all cases, formulations were bioequivalent in terms of AUCi. Differences in Cmax occurred. For Cmax, MODERATE-A and MODERATE-B were bioequivalent. In both study 1 and 2 FAST Cmax was over 10% greater than MODERATE-A Cmax; they were considered modestly dissimilar. FAST Cmax was 20% greater than MODERATE-B Cmax; they were considered markedly dissimilar.

In Fig. 2, there was rank-order agreement in mean Cmax ratio and AUCi ratio across the four comparisons. In the comparison of MODERATE-A and MODERATE-B, Cmax and AUCi ratios were lowest (lower left data point). For FAST vs. MODERATE-B, Cmax and AUCi ratios were highest (upper right data point). A criticism of Cmax and AUCi as bioequivalence metrics is that AUCi and Cmax are each measures of extent of drug availability. Hence, extent is tested twice. Figure 2 supports this criticism, although clearly Cmax and AUCi are not identical. Interestingly, for the two comparisons where Cmax fails but AUCi passes, an upper AUCi criteria of 115% would give concordance between Cmax and AUCi, in terms of passing or failing.

In Table II, skewness of ln(Cmax ratio) and ln(AUCi ratio) are listed. Except the comparison of FAST vs.

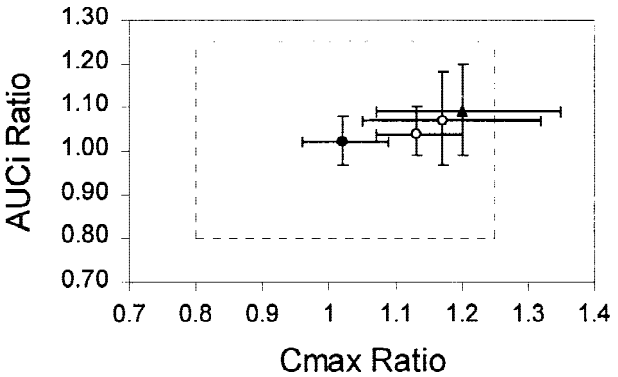

Fig. 2. Relationship between AUCi ratio and Cmax ratio. Error bars denote 90% confidence interval. The box circumscribed by dashed lines is the region of bioequivalence, based upon AUCi (0.8–1.25) and Cmax (0.8–1.25). From lower left to upper right, points are MODERATE-A vs. MODERATE-B (●), FAST vs. MODERATE-A (study 1) (○), FAST vs. MODERATE (study 2) (○), and FAST vs. MODERATE-B (▲). Mean Cmax and AUCi ratios followed a rank-order relationship.

MODERATE-A in study 1, either ln(Cmax ratio) or ln(AUCi ratio) showed large, negative skewness. For MODERATE-A vs. MODERATE-B, ln(Cmax ratio) and ln(AUCi ratio) were both very negatively skewed; $g1$ values were –1.65 and –2.47, respectively, due to a single subject. Similarly, for FAST vs. MODERATE-A in study 2, $g1$ = –1.08 for ln(Cmax ratio) due to a single subject. Since a confidence interval approach assumes a normal distribution, a desirable tendency of a bioequivalence metric is $g1 = 0$. These skewness values deviate substantially from 0, suggesting that Cmax and AUCi are not optimal bioequivalence metrics, from a statistical viewpoint. It should be noted that ln transformation of Cmax and AUCi is not always performed. However, the use of untransformed Cmax and AUCi generally did not improve skewness here.

Novel DCC Metrics: Numerical Values and Comparison to Cmax and AUCi

Table I, the least squares mean (and upper 95% confidence limit) of ρ, ρ_m, δ_a, and δ_s from each comparison is listed. Values for f_2 are not shown since it exhibited poor distribution properties for statistical analysis (see below and Table II). Numerical mean values for ρ, ρ_m, δ_a, and δ_s were about 1.3, 0.25, 0.2, and 0.06, respectively, and reflected the nature of each metric. For ρ, values of about 1.3 show ρ to represent the ratio of the higher concentration to the lower concentration, with approximately a 30% difference. A 30% difference between formulations across the entire plasma profile is common (3). For ρ_m, values of about 0.25 reflect that $\rho_m \cong \rho - 1$. Table I does not indicate $\rho_m = \rho - 1$, but rather $\rho_m < \rho - 1$ (i.e. $0.25 < 1.3 - 1$). In Table I, $\rho_m < \rho - 1$ since exp(mean $\ln(x_i)$) < mean of exp($\ln(x_i)$); the point estimate of ρ_m in Table I is computed from the traditional exp(mean $\ln(x_i)$) method (24). Hence, while a derivation of $\rho_m = \rho - 1$ from Eq. 3a is correct, application of the ln transformed method will show $\rho_m < \rho - 1$, as in Table I.

There was rank-order agreement among mean ρ, ρ_m, δ_a, and δ_s values. The rank-order (according to increasing difference) was:

MODERATE-A vs. MODERATE-B
FAST vs. MODERATE-A (study 2)
FAST vs. MODERATE-A (study 1)
FAST vs. MODERATE-B

The upper 95% confidence interval of ρ, ρ_m, and δ_s (but not δ_a) also exhibited this rank-order. This rank-order did not agree with that for Cmax and AUCi, where FAST vs. MODERATE-A comparisons in studies 1 and 2 were switched (as in Table I).

As illustrated in Fig. 3a–d and listed in Table I, MODERATE-A vs. MODERATE-B gave lowest values for ρ, ρ_m, δ_a, and δ_s, as well as the lowest Cmax and AUCi ratios. In Fig. 3a–d, the lower left data point shows the lowest values to be the DCC metrics and Cmax values for MODERATE-A vs. MODERATE-B. Similarly, FAST vs. MODERATE-B gave highest values for ρ, ρ_m, δ_a, and δ_s, as well as Cmax ratio [and AUCi ratio] (upper right data point).

However, for the two middle profile comparisons involving FAST vs. MODERATE-A, the rank-order for mean DCC metrics differed from the rank-order for Cmax and AUCi. For FAST vs. MODERATE-A, Cmax and AUCi ratios from study 2 exceeded study 1 ratios. Cmax in study 2 failed bioequivalence. However, for ρ, ρ_m, δ_a, and δ_s, study 2 gave smaller mean metric values than study 1. In Fig. 3a–d, for each of the DCC metrics, there was not a rank-order agreement between Cmax and the mean DCC metric.

Moreover, from Fig. 3 and Table I, there was not rank-order agreement between Cmax and the upper 95% confidence limit for ρ, ρ_m, and δ_s. For each the ρ, ρ_m, and δ_s upper confidence limits, FAST vs. MODERATE-A comparisons in studies 1 and 2 were switched, relative to Cmax results. For δ_a, there was rank-order agreement between Cmax and δ_a's upper confidence limit. As discussed below, passing and failing profile similarity was determined from the upper confidence limit. These results indicated that while ρ, ρ_m, δ_a, and δ_s are generally similar to Cmax and AUCi ratios in performance, they tended to differ from Cmax and AUCi ratios in rank-ordering profile similarity. These DCC metrics are simply different than Cmax and AUCi.

This difference between these novel DCC metrics and Cmax and AUCi ratios reflects the distinction between a DCC approach and the approach using Cmax and AUCi to compare profiles. A DCC approach compares profiles at all time points. Arguably, Cmax and AUCi do not compare entire profiles at all points in the same way that ρ, ρ_m, δ_a, and δ_s achieve profile comparison. Cmax perhaps concerns only one time point, and often not the same time point between test and reference. An infinite number of profiles can calculate to the same AUCi. While it may be debatable that metrics such as ρ, ρ_m, δ_a, and δ_s more comprehensively compare entire profiles, or whether more comprehensive comparisons are needed, Fig. 3a–d indicates that these novel DCC metrics perform differently than Cmax and AUCi ratios in profile comparisons. This observation is consistent with previous observations, where DCC metrics better detect curve shifts (i.e. drug absorption lag times) and differences between multiple peak profiles, compared to AUCi and Cmax (3).

Novel DCC Metrics: Distribution Properties

In addition to arguably better profile difference detection, the DCC metrics exhibited better distribution proper-

Table II. Skewness Values (g1) of Bioequivalence Metrics

Formulations	ln(Cmax ratio)	ln(AUCi ratio)	ln(ρ)	ln(ρ_m)	ln(δ_a)	ln(δ_s)	*f2*
MODERATE-A vs. MODERATE-B (study 2)	–1.65	–2.47	1.77	0.549	0.454	0.375[a]	–1.49
FAST vs. MODERATE-A (study 1)	0.498	–0.197	0.309	–0.261[a]	–0.432	–0.446	–0.886
FAST vs. MODERATE-A (study 2)	–1.08	0.116	–0.086[a]	–0.454	–0.463	–0.460	–1.04
FAST vs. MODERATE-B (study 2)	–0.303	–1.60	1.09	0.323	0.576	0.242[a]	–1.54

[a] Least skewed.

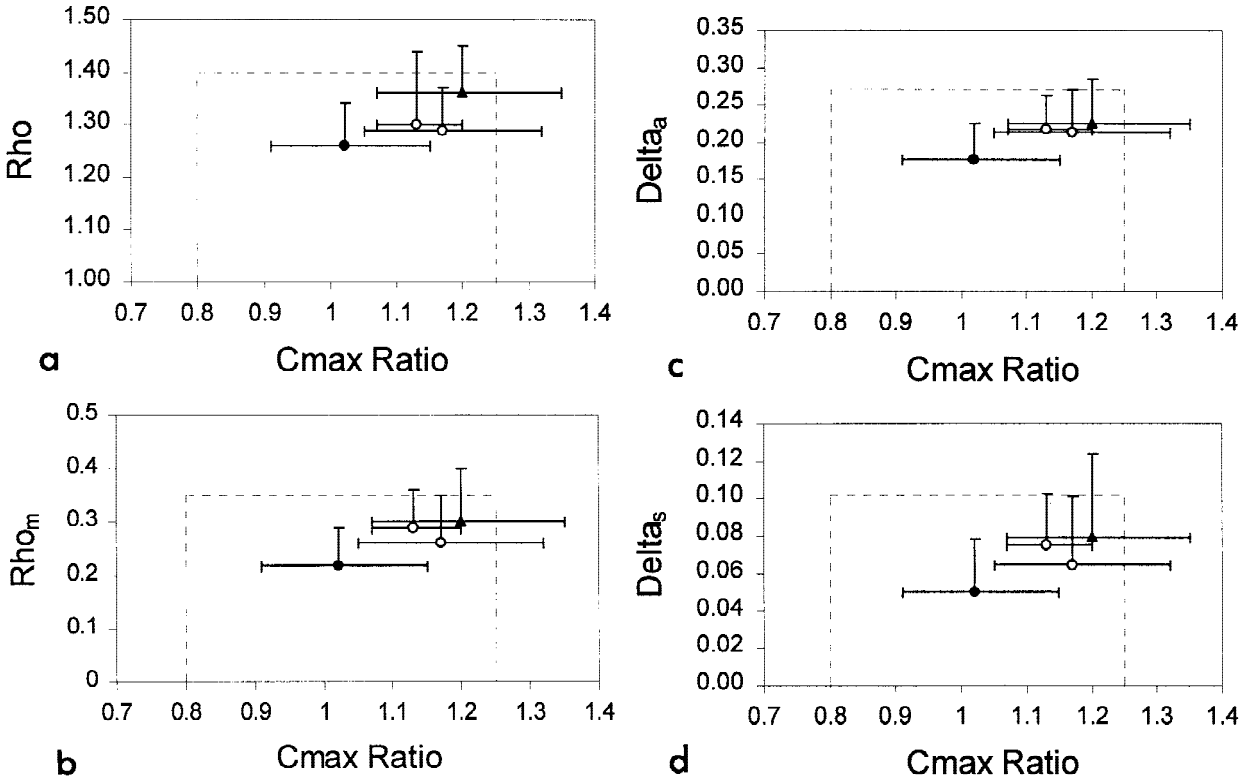

Fig. 3. Relationship between Cmax ratio and the four novel bioequivalence metrics (a) ρ, (b) ρ_m, (c) δ_a, and (d) δ_s. Error bars denote 90% confidence interval for Cmax, and upper 95% confidence limit for novel bioequivalence limit. The region circumscribed by dashed lines is the region of bioequivalence, based upon Cmax (0.8–1.25) and the novel bioequivalence metric ($\rho \leq 1.40$, $\rho_m \leq 0.35$, $\delta_a \leq 0.27$, and $\delta_s \leq 0.102$). The novel metrics generally reflected Cmax ratio, but frequently provided a different rank-order than Cmax ratio for profile similarity. Points are MODERATE-A vs. MODERATE-B (●), FAST vs. MODERATE-A (study 1) (○), FAST vs. MODERATE (study 2) (○), and FAST vs. MODERATE-B (▲).

ties. The 95% confidence limit approach assumes a normal distribution. A practical method to assess this assumption is to measure skewness (26,27). Table I shows that distributions for ln(ρ), ln(ρ_m), ln(δ_a), and ln(δ_s) were more symmetrical than for ln(Cmax ratio) and ln(AUCi ratio). In(ρ_m), ln(δ_a), and ln(δ_s) performed the best. $g1$ for the novel DCC metrics were generally more close to 0, than ln(Cmax ratio) and ln (AUCi ratio). In most cases, $g1$ was ± 0.5 of zero. Meanwhile, except FAST vs. MODERATE-A in study 1, either ln (Cmax ratio) or ln(AUCi ratio) showed large, negative skewness ($g1 < -1$). f_2 exhibited poor distribution properties. f_2 was often strongly negatively skewed ($g1 < -1$). ln(f_2) was even more negatively skewed (data not shown).

For FAST vs. MODERATE-A in study 2, one subject showed a low Cmax ratio of 76.4%, contributing to a negatively skewed ln(Cmax ratio) distribution ($g1 = -1.08$). Distribution skewness of ln(ρ), ln(ρ_m), ln(δ_a), and ln(δ_s) did not suffer from this subject. This subject did not even provide the largest ln(ρ), ln(ρ_m), ln(δ_a), and ln(δ_s) values, but rather only the third or fourth largest values, since all of the other data points are including in the profile comparison, and are not as disparate. Similarly, for MODERATE-A vs. MODERATE-B, ln(Cmax ratio) and ln(AUCi ratio) were both very negatively skewed due to a single subject. The distributions of the novel metrics were better behaved.

Because of its more favorable distribution characteristics and its easy interpretability (i.e. its interpretation as a ratio), ρ_m is the most promising DCC metric. Clearly, however, more experience with this and other DCC metrics is needed.

Novel DCC Metrics: Confidence Limits and Suggested Acceptance Criteria

In considering bioequivalence acceptance limits for ρ, ρ_m, δ_a, and δ_s, Table I lists the upper 95% confidence acceptance limits that provide maximum concordance with Cmax and AUCi bioequivalence results. The limits are: 1.40 for ρ, 0.35 for ρ_m, 0.27 for δ_a, and 0.102 for δ_s. Of course, these identified limits are exploratory. Much more experience is needed to assess the therapeutic validity of such criteria. These values are consistent with the derivation and nature of each metric, and reflect the familiarity of Cmax and AUCi criteria. For example, the criteria $\rho \leq 1.40$ and $\rho_m \leq 0.35$ (rather than $\rho \leq 1.25$ and $\rho_m \leq 0.25$) are consistent with bioequivalent profiles often differing by over 30% on simple average across all time points (3). Also, an evaluation where $\delta_a \leq 0.27$ closely reflects the −80%–+125% range for Cmax and AUCi.

As emphasized above and in Fig. 3a–d, ρ, ρ_m, δ_a, and δ_s are similar to but different from Cmax and AUCi ratios. The broad similarities are evident in Fig. 3a–d and listed in Table I. For the comparison of MODERATE-A vs. MODERATE-B, which were easily bioequivalent, all four DCC metrics passed. Likewise, for FAST vs. MODERATE-B, which easily failed bioequivalence, all four DCC metrics failed. In Table I, the bioequivalence acceptance criteria of the DCC metrics were: $\rho \leq 1.40$, $\rho_m \leq 0.35$, $\delta_a \leq 0.27$, and $\delta_s \leq 0.102$.

However, Table I indicates discordance between the traditional criteria and several of the DCC metrics, in assessing passing and failing of the FAST vs. MODERATE-A com-

parisons. Cmax and three of the DCC metrics (ρ, ρ_m, and δ_s) provided opposite bioequivalence results for FAST vs. MODERATE-A (study 1) and FAST vs. MODERATE-A (study 2), the two moderately dissimilar profiles. In study 1, FAST and MODERATE-A passed bioequivalence using Cmax and AUCi; in study 2, they failed bioequivalence Cmax and AUCi. δ_a also yielded these same results, based upon a confidence limit of $\delta_a \leq 0.27$ (although mean δ_a was greater in study 1 than study 2). Meanwhile, ρ, ρ_m, and δ_s each failed in study 1, but passed in study 2. Clearly, the consideration of alternative bioequivalence metrics affords the possibility that some metrics will yield a "bioinequivalent" result, in spite of a "bioequivalent" result obtained from Cmax and AUCi, and vice versa.

Potential Role for DCC Metrics

The motivation for this introduction of four novel DCC metrics was the desire to compare the shapes of plasma concentration-time profiles, given the viewpoint that Cmax and AUCi do not sufficiently accomplish this comparison in all cases. Although data sets pass the traditional criteria, curve shifts or delays in absorption of a test formulation relative to reference may result in therapeutic inequivalence for some indications (e.g. pain relief, anti-anxiety, nasal decongestant, treatment of diarrhea). DCC metrics have previously been shown to better detect curve shifts (i.e. drug absorption lag times) and differences between multiple peak profiles, than Cmax and AUCi (3).

From the work presented here, it has been shown that DCC metrics can show concordance with the traditional criteria (i.e. traditional criteria and DCC metric both agree and conclude bioequivalence, or both conclude bioinequivalence). It has also been shown that disconcordant results can occur (i.e. traditional criteria passes while DCC fails, and vice versa). All four combinations of concordance and discordance between traditional and DCC metrics was observed here.

Of particular interest, FAST was bioequivalent to MODERATE-A in study 1 in traditional criteria (Fig. 1), yet several of the DCC metrics concluded bioinequivalence, as FAST's profile had risen earlier and declined earlier relative to the profile of MODERATE-A. Meanwhile, in study 2, FAST was bioinequivalent to MODERATE-A in the traditional criteria due to Cmax (Fig. 2), yet several of the DCC metrics concluded bioequivalence, as the rise and fall of the profiles were similar, except in the timeframe around Cmax. These results indicate that DCC metric can potentially supplement, or indeed replace, the existing criteria for bioequivalence. Clearly, however, more experience with DCC metrics is needed.

In summary, the object of this work was to devise four new DCC metrics and examine each metric's distribution properties and performance characteristics. The DCC metrics ρ, ρ_m, δ_a, and δ_s exhibited more favorable distributions that Cmax and AUCi, which were frequently skewed. While it may be debatable that DCC metrics more comprehensively compare entire profiles, or whether more comprehensive comparisons are needed, the DCC metrics performed differently than Cmax and AUCi ratios in profile comparisons. Each DCC metric appears to measure "exposure" rather "rate and extent," and appears to achieve such an assessment in a single valuation. Possible bioequivalence acceptance criteria are: $\rho \leq 1.40$, $\rho_m \leq 0.35$, $\delta_a \leq 0.27$ and $\delta_s \leq 0.102$. Given the potential need for "exposure" metrics, these DCC metrics (particularly ρ_m) are promising bioequivalence metrics.

REFERENCES

1. A. Rescigno. Bioequivalence. *Pharm. Res.* **9**:925–928 (1992).
2. V. M. Chincilli and R. K. Elswick. The multivariate assessment of bioequivalence. *J. Biopharm. Stat.* **7**:113–123 (1997).
3. S. A. Marston and J. E. Polli. Evaluation of direct curve comparison metrics aplied to pharacokinetic profiles and relative bioavailability and bioequivalence. *Pharm. Res.* **14**:1363–1369 (1997).
4. *Approved Drug Products with Therapeutic Equivalence Evaluations,* 17th edition. FDA Center for Drug Evaluation and Research, Rockville, MD, 1997.
5. *Federal Register.* **42**:1648 (1977).
6. F. Y. Bois, T. N. Tozer, W. W. Hauck, M. L. Chen, R. Patnaik, and R. L. Williams. Bioequivalence: Performance of several measures of extent of absorption. *Pharm. Res.* **11**:715–722 (1994).
7. R. P. Basson, B. J. Cerimel, K. A. DeSante, and D. C. Howey. Tmax: An unconfounded metric for rate of absorption in single dose bioquivalence studies. *Pharm. Res.* **13**:324–328 (1996).
8. A. Rostami-Hodjegan, P. R. Jackson, and G. T. Tucker. Sensitivity of indirect metrics for assessing "rate" in bioequivalence studies—"Moving the goalposts" or changing the "game". *J. Pharm. Sci.* **83**:1554–1557 (1994).
9. *Report from Expert Advisory Committee on Biovailability Guidelines and Standards for Bioavailability. Part A. Oral Dosage Forms Used for Systemic Effects,* Health Protection Branch, 1992.
10. R. Schall and H. G. Luus. Comparison of absorption rates in bioquivalence studies of immediate release drug formulations. *Int. J. Clin. Ther. Toxicol.* **30**:153–159 (1992).
11. L. F. Lacey O. N. Keene, C. Duquesnoy, and A. Bye. Evaluation of different measures of rate of drug absorption in comparative pharmacokinetic studies. *J. Pharm. Sci.* **83**:212–215 (1994).
12. L. Tothfalusi and L. Endrenyi. Without extrapolation, Cmax/AUC is an effective metric in investigations of bioequivalence. *Pharm. Res.* **12**:937–942 (1995).
13. T. N. Tozer, F. Y. Bois, W. W. Hauck, M. L. Chen, and R. L. Williams. Absorption rate vs. exposure: Which is more useful for bioequivalence testing? *Pharm. Res.* **13**:453–456 (1996).
14. M. Olling, T. T. Mensinga, D. M. Barends, C. Groen, O. A. Lake, and J. Meulenbet. Bioavailability of carbamazepine from four different products and the occurrence of side effects. *Biopharm. Drug Dispos.* **20**:19–28 (1999).
15. M. L. Chen. An alternative approach for assessment of rate of absorption in bioequivalence studies. *Pharm. Res.* **9**:1380–1385 (1992).
16. R. P. Basson, A. Ghosh, B. J. Cerimele, K. A. DeSante, and D. C. Howey. Why rate of absorption inferences in single dose bioequivalence studies are often inappropriate. *Pharm. Res.* **15**:276–279 (1998).
17. L. Endrenyi, F. Csizmadia, L. Tothfalusi, and M. L. Chen. Metrics comparing simulated early concentration profiles for the determination of bioequivalence. *Pharm. Res.* **15**:1292–1299 (1998).
18. L. Endrenyi and L. Tothalusi. Secondary metrics for the assessment of bioequivalence. *J. Pharm. Sci.* **86**:401–402 (1997).
19. L. Endrenyi and L. Tothalusi. Response to "Is Cmax/AUC useful for bioequivalence testing?" *J. Pharm. Sci.* **86**:1505–1506 (1997).
20. A. Rostami-Hodjega and G. T. Tucher. Is Cmax/AUC useful for bioequivalence testing? *J. Pharm. Sci.* **86**:1504–1505 (1997).
21. T. N. Tozer and W. W. Hauck. Cmax/AUC, a commentary. *Pharm. Res.* **14**:967–968 (1997).

22. T. N. Tozer, F. Y. Bois, W. W. Hauck, M. L. Chen, and R. L. Williams. Absorption rate vs. exposure: Which is more useful for bioequivalence testing? *Pharm. Res.* **13**:453–456 (1996).
23. W. R. Garnett, B. Levy, A. M. McLean, Y. Zhang, R. A. Couch, E. M. Rudnic, J. M. Pellock, and G. W. Belendiuk. Pharmacokinetic evaluation of a twice-daily extended-release carbamazepine (CBZ) and four-times-daily immediate-release CBZ in patients with epilepsy. *Epilepsia* **39**:274–279 (1998).
24. *Statistical Procedures for Bioequivalence Studies Using a Standard Two-treatment Crossover Design. Center for Drug Evaluation and Research*, FDA, Rockville, MD, 1992.
25. J. E. Polli, G. S. Rekhi, L. L. Augsburger, and V. P. Shah. Methods to compare dissoution profiles anda rationale for wide dissolution specifications for metoprolol tartrate tablets. *J. Pharm. Sci.* **86**:690–700 (1997).
26. G. W. Snedecor and W. G. Cochran. *Statistical Methods*, The Iowa State University Press, Ames, IA, 1980.
27. L. Wilkinson, G. Blank, and C. Gruber. *Desktop Data Analysis with SYSTAT*, Prentice Hall, Upper Saddle River, NJ, 1996.