Bioequivalence: Performance of Several Measures of Rate of Absorption

Frédéric Y. Bois,[1] Thomas N. Tozer,[2] Walter W. Hauck,[3] Mei-Ling Chen,[4] Rabindra Patnaik,[4] and Roger L. Williams[4]

Received October 26, 1993; accepted February 18, 1994.

The highest point of the plasma concentration-time profile, C_{max}, is currently used by regulatory agencies to assess the rate of drug absorption after single dose administration of oral products. It is, however, quite insensitive, and a number of new measures of rate have been proposed. Using simulations, several approaches toward measuring rate were tested. A set of model scenarios for drugs with typical mean characteristics and statistical distributions was investigated. Using different kinetic models of disposition, the time course of the concentration in plasma was simulated. Intraindividual and interindividual variability and assay error were modeled using Monte Carlo techniques. The accuracy, precision, and ease of use of the various measures of rate were evaluated by simulating cross-over design clinical trials and then determining the probability of declaring bioequivalence as a function of differences in rates of absorption between test and reference formulations. All of the rate measures tested showed a degree of insensitivity to changes in rate and no universally superior measure was found. Indeed, the main conclusion is that the choice of a measure should be based on simulations of the particular situation in a bioequivalence trials.

INTRODUCTION

The most common method for assessing rate of drug absorption after a single dose administration of an oral product is to record C_{max}, the highest drug concentration achieved in the plasma for a given individual. However, C_{max} depends on both rate and extent of absorption, and is a particularly good measure of the latter, as previously shown (1). It is, however, quite insensitive to changes in rate. The time of observation of C_{max}, t_{max}, is widely recognized as having reliability problems, particularly because well-established statistical tests applicable to such discrete variables (2) are lacking. The partial area under the curve, AUC_p computed up to t_{max}, has recently been proposed as a replacement of C_{max} (3). Additional measures have been proposed in which C_{max} is divided by various factors to free it from dependence on extent. These factors include: The AUC from zero to infinity, leading to C_{max}/AUC_∞ (4); t_{max}, leading to C_{max}/t_{max}; or the AUC up to t_{max}, giving $C_{max}/AUC_{t_{max}}$. Other techniques, feathered slope (SL_f) and feathered AUC (AUC_f) derived from peeling (5), may be used to subtract an estimated elimination component from the concentration curve to isolate drug absorption. Finally, the first few points of the concentration-time curve, or the AUC up to those points (for improved stability) were tested, alone or in combination, as measures of rate of absorption.

We examined, through simulations, the potential of the measures to assess bioequivalence. The conditions of a typical clinical bioequivalence trial were simulated in eight scenarios of absorption and disposition kinetics. The scenarios incorporate the following: ratios of absorption and elimination rate constants of 0.25 and 4, zero-order and first-order absorption kinetics, limits of quantification of 1% and 10% of the mean peak concentration, presence of a lag-time, highly variable first-pass elimination, and two-compartment distribution characteristics. The time course of the concentration in the plasma was simulated, with intraindividual and interindividual variability and assay error modeled using Monte Carlo techniques. For each scenario, the rate of absorption was assessed by the different measures investigated.

Computer simulations allow control of the model parameters while conducting a typical bioequivalence trial. Known differences between test and reference formulations and known levels of variability were introduced. The ability of given measures to uncover the true underlying differences in rate were thereby assessed. The reliability of selected measures was evaluated by power analyses (to assess the frequency of trials in which bioequivalence is declared).

METHODS

Simulation Framework

To evaluate the various measures of rate of absorption, it was necessary to simulate the variability typically encountered in humans. To accomplish this, Monte Carlo simulations were used to generate data sets of clinical trials to which the measures were applied. Details are given in the previous report on extent (1). A summary is given here.

In all cases, 1540 clinical trials were simulated. Each trial was a cross-over design with 24 subjects and two drug formulations (test, T, and reference, R). A random sequence effect, with average zero, was introduced by segregating the subjects into two groups of 12.

A standard statistical model of errors and variabilities in population pharmacokinetics (6) was used to simulate the trials. From population distributions, a set of pharmacokinetic parameter values was sampled for each subject. At each trial period, intraindividual variability was added to the subject's baseline values, forming new parameter values These values were assumed to remain constant over a trial period. Two periods were simulated, during which the two formulations were administered. The difference between the two formulations was introduced by changing the value of the rate of absorption for each subject's test period. These changes amounted to a fixed fraction of the mean population value, and were the same for each individual. Assay error was added to the plasma drug concentrations simulated by the model. The various variabilities were taken to be the same for the two formulations.

[1] Indoor Environment Program, Mail Stop 90-3058, Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, CA 94720
[2] School of Pharmacy, University of California San Francisco, San Francisco, CA 94143-0446.
[3] Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA 94143-0560. Currently: Dept. of Clinical Pharmacology, Jefferson Medical College, Philadelphia, PA.
[4] Department of Health and Human Services, Food and Drug Administration., Center for Drug Evaluation and Research, Office of Generic Drugs, 7500 Standish Place (HFD-600), Rockville, MD 20855-2737

0724-8741/94/0700-0966$07.00/0

Individual plasma concentration values at the defined sampling times were simulated using specific interindividual and intraindividual distributions for each parameter, as described above (Tables I to IV). A body weight of 70 kg was assumed in all cases.

To avoid unrealistic values all normal distributions were truncated. Analytical assay errors were generated from truncated normal distributions with mean zero, *CV* of 10%, truncation at ± 4 *CV* plus a fixed term equal to the product of the assay *CV* and the limit of quantification, *LQ*, defined as a fraction of the theoretical C_{peak} computed with the mean interindividual parameter values (Table I). Concentrations values lower than *LQ* were treated as null values in the subsequent computations.

Scenario Definitions

The pharmacokinetic models and parameter distributions used were grouped in eight scenarios which reflect situations commonly encountered, or of special interest, when testing bioequivalence. These scenarios have been previously described in detail (1).

Baseline Simulation Scenario

The baseline scenario includes: one-compartment distribution kinetics; first-order absorption and limitations; ratio of absorption to elimination rate constants of 4; no lag-time; and an *LQ* equal to 1% of the theoretical mean peak concentration, C_{peak}, for the reference formulation after oral administration.

For all studies an arbitrary oral bolus dose of 500 mg was used. Simulated sampling times were 0, 0.25, 0.5, 1, 1.5, 2 hr and then every 2 hours up to 16 hours.

Alternative Scenarios

A set of alternative scenarios were examined. In each case only differences from the baseline scenario are described.

Low Sensitivity This scenario corresponds to a situation in which the concentration cannot be followed over a wide range. The *LQ* was set at 10% of the mean interindividual C_{peak} for the reference formulation.

Zero-order Absorption The duration of input was set to $2/k_a$, where k_a is the absorption rate constant in the baseline scenario (Table I). The coefficient of variation of the duration of input had the same *CV* and truncation as the baseline k_a.

Presence of a Lag-Time A random lag-time of an hour, on average, was introduced for both the test and the reference formulations. A normal distribution with a *CV* of 50%, truncated to ± 2 *CV* was used.

Low Absorption/Elimination Ratio ("Flip-Flop") The ratio of absorption/elimination rate constants was fixed at 0.25. The simulated sampling lines were then: 0, 1, 2, 4, 6, 8, 12, 16, 20, 24, 32, 40 and 48 hours.

Low F This scenario corresponds to the situation of high first-pass elimination. The extent of absorption, *F*, was sampled following a uniform distribution with range of 0.05 to 0.15 (mean 0.1) for interindividual variability, and a range of −0.05 to +0.05 for intraindividual variability. *LQ* was maintained at the same value as in the baseline case.

Two-compartment Models Two two-compartment models with first order input into, and elimination from, the central compartment were studied. In the first model the elimination/distribution ratio of the rate constants k_{10}/k_{21} was fixed at 2.5 (model I) and in the second set at 0.4 (model II), k_{10} being equal to *CL/V*. The ratio k_{12}/k_{21} was 4 in both cases. Tables III and IV give the interindividual and intraindividual distributions used for these models. For each model two cases, *LQ* at 1% and at 10% of the reference C_{peak}, were investigated.

Measures of Rate

The following measures of rate were evaluated:

C_{max}, t_{max}

C_{max} was determined by finding the highest recorded plasma concentration for a given individual and a given trial period. The corresponding time is t_{max}, obtained without interpolation.

Partial AUC (AUC_p)

In the context of a cross-over bioequivalence trial, as implemented in our simulations, each individual received the reference and the test formulations (in two separate administrations). Following Chen (3), the area under the time-concentration curve up to t_{max} of the reference formulation, in a given individual, was computed for both formulations by the trapezoidal rule without transformation or extrapolation.

C_{max}/AUC_∞

C_{max} was computed as given above and AUC_∞ was computed as the sum of the areas up to the last observable point (obtained using the trapezoidal method) and remaining (estimated from a simple exponential passing through the least-square estimate of the last observable data point). The rate

Table I. Distribution type, mean, coefficient of variation (*CV*) and truncation for interindividual parameters in the one-compartment baseline scenario.

Parameter	Distribution	Mean	CV (%)	Truncation
Volume of distribution, *V*	normal	1 L/kg	10	±3 *SD*
Clearance, *CL*	normal	0.347 L/(hr × kg)	20	±3 *SD*
Absorption rate constant, k_a	normal	1.39 hr^{-1}	20	±3 *SD*
Bioavailability, *F*	uniform	0.5	11.5[a]	0.4–0.6

[a] This value is equal to $100 \times (0.6 - 0.4)/(0.5 \times \sqrt{12})$ since the distribution is uniform.

Table II. Distribution type, mean, coefficient of variation (*CV*) and truncation for intraindividual parameters in the one-compartment baseline scenario.

Parameter	Distribution	Mean	*CV* (%)	Truncation
Volume of distribution, V	normal	vps[a]	10	±3 *SD*
Clearance, CL	normal	vps	20	±3 *SD*
absorption rate constant, k_a	normal	vps	20	±3 *SD*
Bioavailability, F	uniform	vps	—[b]	−0.1 to 0.1

[a] vps = parameter value previously sampled (for a given individual) (see text).
[b] This value is equal to $100 \times 2/(\text{vps} \times \sqrt{12})$ since the distribution is uniform.

constant of the exponential was obtained by fitting a straight line to the last four observable data points, after log-transformation. AUC_∞ was denoted AUC_{inf4} in our previous paper (1).

C_{max}/t_{max}

This measure was simply the ratio of the observed C_{max} and the time of its occurrence.

C_{max}/AUC_{tmax}

This measure was the ratio of C_{max} and AUC, computed by the trapezoidal rule up to t_{max} for a given individual and period.

Feathered Slope (SL_f)

A straight line was fitted by least squares to the last four observable points (after log-transformation). Concentrations values along that line were computed by extrapolation for the first four time points at which measurements were made. The differences between the extrapolated and measured concentrations were refitted (in natural log space) to a straight-line, called the feathered curve. The negative of the slope of the feathered curve is the absorption rate constant. The method is based on a first-order absorption model with the absorption process being faster than that of elimination.

Feathered AUC (AUC_f)

This method is similar to the previous except that the area under the feathered curve, rather than its slope, was computed by the trapezoidal method.

Cumulative AUCs (AUC_1, AUC_2, AUC_3)

The areas under the time-concentration curve were directly computed by the trapezoidal method from the simulation output up to each of three times, namely 1/4, 1/2 and 1 times the average t_{peak} for the reference formulation. In the simulations these times were the first, second, and fourth times of observation for each scenario.

Concentration Values (C_1, C_2, C_3)

The plasma concentration values were directly obtained from the simulation output, up to the same three times as above.

Reliability Analysis

Two sets of simulations (1540 trials each) were performed for each scenario. In the first, the same drug formulation was readministered to each individual. This set of simulations provides a "null" distribution for the difference in the rate of absorption between two formulations, and test the ability of the various measures to show bioequivalence.

In the second, the administration of two drug formulations, differing in rate by 25% ($k_{a\ test}/k_{a\ reference} = 1.25$) was simulated. This difference corresponds to the statistical null hypothesis for bioequivalence since measures of rate for the two formulations are considered equivalent if within 80% and 125%. The ability of the various measures to show the 25% difference between formulations was then evaluated using this simulation set.

Power Analysis

Sets of simulations were performed to determine the

Table III. Distribution type, mean, coefficient of variation (*CV*) and truncation for interindividual parameters in the two-compartment models I and II.

Parameter	Distribution	Mean		*CV* (%)	Truncation
Volume of distribution, V_1	normal	1 L/kg		10	±3 *SD*
Clearance, CL	normal	0.5 L/(hr × kg)		20	±3 *SD*
Absorption rate constant, k_a	normal	2.0 hr^{-1}		20	±3 *SD*
Central to peripheral distribution rate constant, k_{12}	normal	0.2 hr^{-1} 1.25 hr^{-1}	(model I) (model II)	20	±3 *SD*
Peripheral to central distribution rate constant, k_{21}	normal	0.05 hr^{-1} 0.3125 hr^{-1}	(model I) (model II)	20	±3 *SD*
Bioavailability, F	uniform	0.5		11.5[a]	0.4–0.6

[a] This value is equal to $100 \times (0.6 - 0.4)/(0.5 \times \sqrt{12})$ since the distribution is uniform.

Table IV. Distribution type, mean, coefficient of variation (*CV*) and truncation for intraindividual parameters in the two-compartment models I and II.

Parameter	Distribution	Mean	*CV* (%)	Truncation
Volume of distribution, V_1	normal	vps[a]	10	±3 *SD*
Clearance, *CL*	normal	vps	20	±3 *SD*
Absorption rate constant, k_a	normal	vps	20	±3 *SD*
Central to peripheral distribution rate constant, k_{12}	normal	vps	20	±3 *SD*
Peripheral to central distribution rate constant, k_{21}	normal	vps	20	±3 *SD*
Bioavailability, *F*	uniform	vps	—[b]	−0.1 to 0.1

[a] vps = parameter value previously sampled (for a given individual) (see text).
[b] This value is equal to $100 \times 0.2/(\text{vps} \times \sqrt{12})$ since the distribution is uniform.

relationship between the statistical power of the procedures examined (i.e., the probability of rejecting the null hypothesis of bioequivalence in a clinical trial) and the difference in rate between the test and reference formulations. For each of a series of values of test/reference absorption rate ratios, 1540 clinical trials were generated by Monte Carlo simulations. In each trial the 24 subjects were randomly segregated into two groups to simulate sequence effects and two one-sided *t*-tests procedure (90% confidence interval) were performed after log-transformation of the measures of rate (7, 8). In the rare case of missing data (e.g., all zero values for concentrations leading to an unobservable C_{max}) the corresponding individual was dropped from the trial. The fraction of trials in which bioequivalence was declared was recorded. This fraction corresponds to the probability of declaring bioequivalence given a typical design of a clinical trial, human and analytical variability, data treatment procedures and statistical analyses.

With no difference between the test and reference formulations one would want to conclude bioequivalence in 100% of the trials. In contrast, with a 25% difference in rate, a value currently used in regulatory practice, bioequivalence should be declared in no more than 5% of the trials (i.e., a 5% consumer risk). The ability to meet these criteria depends on the quality of the measure of rate of absorption used.

The results are presented in the form of power curves, in which the x-axis is the test/reference absorption rate ratio used in the simulations. The y-axis is the corresponding probability of declaring bioequivalence, when assessed by two one-sided *t*-tests applied to a given measure of rate. Each point was obtained from 1540 simulated trials. For display, all the curves were smoothed by spline interpolation to reduce jaggedness due to statistical variability in our power estimates.

RESULTS

Reliability Analysis

Table V gives the mean and standard deviation (SD) of the percent differences of the mean measures in the test and reference groups $100\times(T\text{-}R)/R$, over 1540 simulated clinical trials of size $N = 24$. The test and reference formulations have the same rate and extent of absorption and so the relative differences should be centered on zero (i.e., clinical trials should on average indicate no difference). The reliability of the various measures can be compared by examining Table V column by column. C_{max} has in general a low bias: 0.3% at most. Associated SDs are at most 7%, which means that the large majority of clinical trials, analyzed using C_{max}, should correctly identify the two formulations as equivalent. The t_{max} measure also has a low bias but the associated SDs are systematically larger than those of C_{max}. AUC_p behaves very similarly to C_{max} (low bias, low SD) except in the case of random lag time in which its SD increases markedly. C_{max}/AUC_∞ behaves also like C_{max}, except in the case of two-compartment kinetics (low *F*) where its SD is quite high and the mean quite far from zero. C_{max}/t_{max} and C_{max}/AUC_{tmax} both have low bias but very large SDs, and the same holds for SL_f and AUC_f. For the last two measures, SDs of even 1000% can be found when two-compartment distribution underlies the drug kinetics. The cumulative AUCs (AUC_1, AUC_2, AUC_3) and the concentration values (C_1, C_2, C_3) perform quite well. The SDs of those measures decreases with times. This is expected since the relative experimental error decreases with concentration values increase, and therefore with time, as long as T_{peak} has not been exceeded). These measures are strongly affected by the presence of a random lag-time, in which case they behave very poorly (many zero values which are unusable because of the ratio).

Table VI presents similar data for the case in which the rate of absorption of the underlying models was systematically 25% higher for the test than for the reference formulation. If the measures were proportional to the underlying constant rate k_a, a 25% relative differences would be expected. Practically no measure reaches such a value in any of the cases tested; much smaller changes in the measures themselves were observed. For example, C_{max} differences were generally only of 4% to 7%, except in the case of "flip-flop" for which the average relative difference reached 17% and in the two-compartment scenario II, where it reached 11%. The associated SDs were low (4% to 7%). The t_{max} measure is usually more sensitive, but the associated SDs are higher than those of C_{max}. AUC_p has a high sensitivity overall and potentially the best behavior of all measures, except in the case of random lag-time where a high variability (13%) occurs. C_{max}/AUC_∞ has a behavior similar to that of C_{max}, except in the two-compartment scenarios, where the AUC denominator is highly variable (1). C_{max}/t_{max} and C_{max}/AUC_{tmax} have a good sensitivity but the variabilities of their components cumulate and therefore they have very high SDs. The same is observed for SL_f and AUC_f. The

Table V. Mean ± standard deviation of the percent difference between the average of each rate measure for test and reference groups, when formulations are bioequivalent. Percent differences were obtained from 1540 simulated clinical trials[a].

	Measure													
Scenario	C_{max}	t_{max}	AUC_p	C_{max}/AUC_∞	C_{max}/t_{max}	$C_{max}/AUCt_{max}$	SL_f	AUC_f	AUC_1	AUC_2	AUC_3	C_1	C_2	C_3
Baseline	−0.104	−0.024	0.263	0.156	0.505	0.978	1.11	1.50	−0.032	−0.064	−0.085	−0.032	−0.050	−0.037
	±4.24	±7.96	±4.69	±4.99	±10.2	±11.7	±19.8	±19.7	±6.53	±5.78	±4.47	±6.53	±5.69	±4.56
High *LQ*	0.160	0.199	0.536	0.0943	0.561	0.713	0.503	0.837	0.360	0.298	0.170	0.360	0.269	0.144
	±4.10	±7.78	±4.59	±4.75	±10.0	±11.3	±15.0	±14.8	±6.39	±5.74	±4.42	±6.39	±5.74	±4.57
Zero-Order	0.0801	0.179	1.16	0.234	0.363	0.420	3.89	2.20	0.302	0.239	0.110	0.302	0.225	0.0517
Absorption	±4.34	±6.17	±5.48	±5.38	±9.62	±9.52	±27.5	±23.0	±8.06	±7.68	±5.53	±8.06	±7.96	±5.15
Lag-Time	0.0217	0.432	4.98	0.226	0.671	1.57	2.06	0.820	140	170	4.07	140	110	1.60
	±4.89	±10.8	±12.9	±5.61	±12.0	±16.2	±21.7	±13.9	±1540	±1800	±30.5	±1540	±1160	±18.9
"Flip-Flop"	0.0730	0.0797	−0.716	−0.133	0.760	1.18	1.50	13.6	0.100	0.0872	0.0570	0.100	0.125	0.134
	±6.75	±8.50	±6.97	±4.94	±10.8	±12.1	±18.3	±77.1	±7.14	±6.75	±6.37	±7.14	±7.03	±7.14
Low *F*	0.234	0.124	0.613	0.0431	0.653	0.614	1.10	−8.16	0.522	0.417	0.314	0.522	0.343	0.272
	±3.97	±7.65	±4.53	±5.31	±9.88	±11.1	±16.5	±870	±6.58	±5.77	±4.30	±6.58	±5.74	±4.51
Two-compartment I (elim. > distr.)	0.0220	0.541	0.818	0.758	0.481	0.688	−43.2	−5.76	0.152	0.120	0.0563	0.152	0.127	−0.022
	±4.40	±9.60	±5.02	±9.60	±12.6	±14.5	±1008	±1045	±6.24	±5.58	±4.18	±6.24	±5.57	±4.87
Two-compartment II (elim. < distr.)	0.309	0.709	0.869	0.568	0.836	0.893	−10.9	2.11	0.370	0.357	0.243	0.370	0.393	0.205
	±5.27	±10.4	±5.93	±9.23	±11.0	±13.5	±50.8	±125	±6.13	±5.52	±4.74	±6.13	±5.61	±6.54

(a) For a given measure the percent difference between test and reference is 100 × (Xt − Xref)/Xref, *X* being the average value of the measure of rate of absorption across individuals.

cumulative AUCs and concentrations behave again very poorly in the presence of random lag-time, otherwise AUC_3 and C_3 are the least sensitive while the other two behave quite similarly.

Power Analysis

Power analyses are presented for all scenarios with C_{max}, t_{max}, AUC_p, and C_{max}/AUC_∞, the measures consistently exhibiting the best performances in the reliability studies. Combinations of the cumulative AUCs, or of the concentrations themselves could have been tested but the development of a multivariate statistical test (different from the two one-sided *t*-tests) would have been necessary. The power curves were computed only for k_{aT}/k_{aR} ratios greater than 1. On inspecting Figures 1 and 2, it should be kept in mind that a good measure should have low producer risk. The producer risk is equal to 1 minus the probability of declaring bioequivalence when the two formulations are in fact bioequivalent. One expects the power curve to cross the y-axis at a value close to 1. The consumer risk is the probability of declaring bioequivalence when the formulations are not equivalent, i.e., at a k_a ratio of 1.25 or above. This risk is the only one regulated. A nominal level of 5% (probability level of the *t*-test if its assumptions are satisfied) is used for regulation.

In the case of the baseline scenario (Figure 1A) AUC_p and t_{max} have the best behaviors, with opposing limitations. In the case of bioequivalence ($k_{a\ test}/k_{a\ reference} = 1$) two one-sided *t*-tests applied to AUC_p data leads almost always to the correct conclusion: the producer risk is nearly zero. t_{max}, on the other hand, presents a significant producer risk (about 30%). The consumer risks are far from 5% for a k_a ratio of 1.25. The 5% level is reached only at a k_a ratio of 1.5 when using t_{max} and 1.7 when using AUC_p. The consumer risk is therefore lower for t_{max} than for AUC_p. C_{max} and C_{max}/AUC_∞ have less satisfying behaviors: they lead to a low producer risk, but are insensitive to changes in rate. It is only when k_a for the test is 2.7 times greater than that for the reference that the consumer risk reaches 5%. The same results are obtained in the case of low assay sensitivity and low bioavailability (data not shown).

In presence of zero-order absorption (Figure 1B), as compared to the baseline scenario, t_{max} has improved in terms of producer risk (now at about 10%) and is better overall than AUC_p whose power curve is not much affected by the change in input function. In comparison, C_{max} and C_{max}/AUC_∞ now have even lower sensitivity than in the baseline case. Even with a k_a ratio of 10, C_{max} still leads to a 30% chance of declaring bioequivalence.

When a random lag-time is introduced in both reference and test plasma concentration curves, C_{max} retains very similar power curves, with low producer risk, and quite insensitive to changes in rate (Figure 1C). Its behavior is slightly better than in the baseline case, possibly because of the shift in times caused by the lag. t_{max} has a high producer risk (about 60%) while AUC_p has zero power consistently. AUC_p will never identify the two formulations as bioequivalent, in presence of random lag. One might consider this as an advantage, from the consumer point of view. However the question arises whether the reference formulation itself should have been marketed. If yes (i.e., if despite the lag it is useful and acceptable) it seems unfair to penalize the test formulation for the same behavior.

Table VI. Mean ± standard deviation of the percent difference between the average of each rate measure for test and reference groups, when test and reference differ in k_a by 25%. Percent differences were obtained from 1540 simulated clinical trials[a].

	Measure													
Scenario	C_{max}	t_{max}	AUC_p	C_{max}/AUC_∞	C_{max}/t_{max}	$C_{max}/AUCt_{max}$	SL_f	AUC_f	AUC_1	AUC_2	AUC_3	C_1	C_2	C_3
Baseline	6.47	−9.86	12.0	6.29	19.3	11.6	10.5	−28.0	20.7	18.8	11.6	20.7	16.5	5.29
	±4.28	±7.06	±5.03	±5.44	±12.2	±13.3	±24.0	±14.9	±7.77	±6.74	±4.81	±7.77	±6.50	±4.77
High *LQ*	6.25	−9.92	11.6	6.25	19.1	11.7	10.4	−21.2	19.7	18.0	11.2	19.7	15.9	5.06
	±4.61	±7.23	±5.18	±5.17	±12.5	±13.3	±17.7	±13.0	±7.56	±6.65	±5.05	±7.56	±6.51	±5.00
Zero-Order	4.36	−15.0	15.8	5.03	31.1	23.5	−0.061	−32.1	33.0	31.4	15.1	33.0	29.8	0.712
Absorption	±4.49	±5.80	±5.92	±5.76	±15.2	±14.2	±31.5	±19.0	±12.9	±11.6	±6.12	±12.9	±11.3	±5.03
Lag-Time	7.17	−8.79	15.3	7.18	16.8	16.3	14.5	−6.58	204	406	20.3	204	249	15.5
	±5.27	±10.1	±13.4	±6.08	±13.6	±18.5	±24.5	±13.5	±1500	±7770	±33.8	±1500	±4110	±20.9
"Flip-Flop"	17.1	−7.73	19.3	16.8	26.9	8.92	4.64	79.0	24.4	23.7	20.3	24.4	22.9	16.7
	±7.59	±7.64	±7.93	±5.54	±13.4	±13.0	±18.5	±84.8	±8.41	±7.82	±7.29	±8.41	±8.09	±8.13
Low *F*	6.34	−10.0	11.8	6.52	19.7	12.1	12.9	−33.7	20.3	18.5	11.4	20.3	16.3	5.11
	±4.24	±7.22	±4.80	±5.42	±12.9	±13.8	±19.7	±978	±7.30	±6.42	±4.69	±7.30	±6.45	±4.83
Two-compartment I (elim > distr.)	7.41	−14.4	12.6	6.67	26.2	20.1	−34.0	21.7	18.6	16.2	8.13	18.6	13.0	−0.262
	±4.65	±8.65	±5.33	±10.3	±15.4	±17.1	±101	±523	±7.23	±6.25	±4.41	±7.23	±6.01	±4.94
Two-compartment II (elim < distr.)	11.1	−14.0	15.1	11.7	26.1	17.2	−5.87	−0.204	18.5	16.0	7.44	18.5	12.2	−3.77
	±5.55	±8.39	±6.26	±10.4	±13.5	±15.6	±48.4	±17.8	±7.07	±6.14	±4.96	±7.07	±6.02	±6.16

(a) For a given measure the percent difference between test and reference is 100 × (Xt − Xref)/Xref, X being the average value of the measure of rate of absorption across individuals.

When elimination is faster than absorption, "flip-flop" scenario, the behavior of most measures improves (Figure 1D). Only t_{max} has approximately the same behavior as in the baseline case (with a 30% producer risk). AUC_p has a slightly better behavior than C_{max} (lower producer risk and higher sensitivity). C_{max} and C_{max}/AUC_∞ are much more sensitive than in the baseline case. The nominal power of the t-test (5%) is reached for a k_a ratio of 1.4.

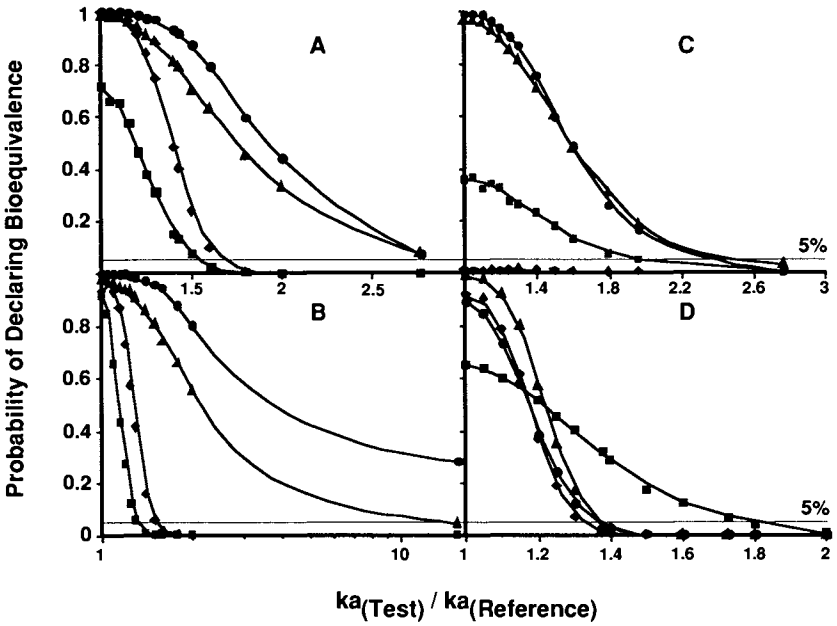

Figure 1. Statistical power curves for four measures of rate of absorption: C_{max} (circles), AUC_p (diamonds), t_{max} (squares), and $C_{max}/AUC\infty$ (triangles). Scenarios: baseline (A), zero-order absorption (B), random lag-time (C), "flip-flop" (D). The probability of declaring bioequivalence using two one-sided t-tests is given as a function of the ratio of absorption rate constants for Test and Reference formulations.

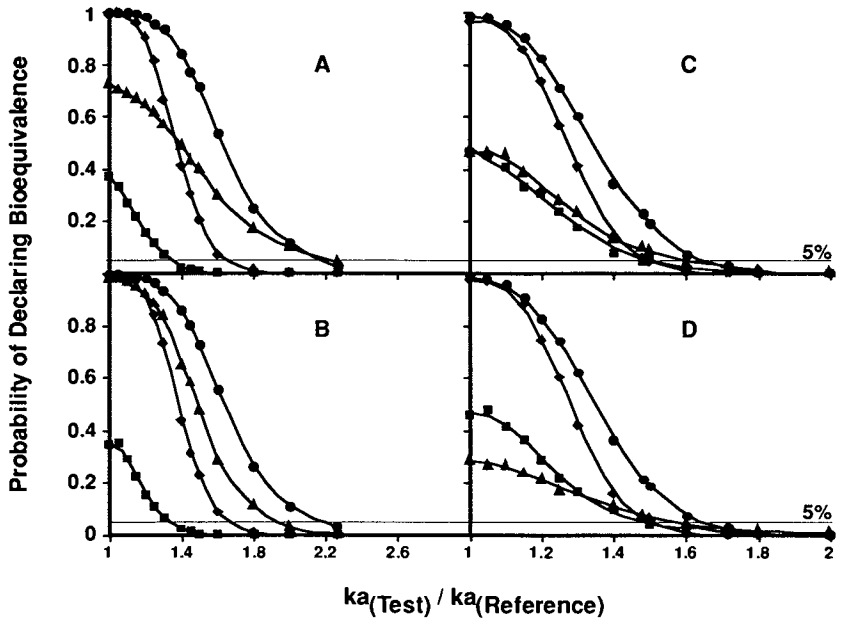

Figure 2. Statistical power curves for four measures of rate of absorption: C_{max} (circles), AUC_p (diamonds), t_{max} (squares), and C_{max}/AUC_∞ (triangles). Scenarios: two-compartment model I (A: with 1% LQ; B: with 10% LQ) and two-compartment model II (C: with 1% LQ; D: with 10% LQ). The probability of declaring bioequivalence using two one-sided t-tests is given as a function of the ratio of absorption rate constants for Test and Reference formulations.

In the case of the two-compartment scenario I (Figure 2A) the four measures studied can easily be ranked. AUC_p has the best behavior: low producer risk and relatively high sensitivity. C_{max}, somewhat less sensitive, is next followed by C_{max}/AUC_∞ which presents a high producer risk. t_{max} is quite sensitive but has a very high producer risk (60%). When the same two-compartment model is coupled to a LQ at 10% of C_{peak} (instead of 1%) the behavior of C_{max}/AUC_∞ improves markedly (Figure 2B). This is due to the improvement of the AUC component whose behavior improves in stability in this case (1). The other measures are not affected by the change in assay sensitivity.

Changes in the structure of the two-compartment model underlying the data can modify considerably the power of the measures studied (Figure 2C and 2D). They all have increased sensitivity in the case of two-compartment scenario II. The 5% nominal level is reached by C_{max} at a k_a ratio of 1.6 instead of 2.2 for model I. AUC_p still has a better behavior (low producer risk, higher sensitivity) than C_{max} at both values of LQ. The power of t_{max} is not much affected by LQ either, or even by the type of two-compartment kinetics driving the plasma concentration curve. C_{max}/AUC_∞ has a much worse behavior for model II than for model I. A lower assay sensitivity markedly decreases its reliability.

DISCUSSION

Assessment of rate and extent, as means of documenting formulation bioequivalence may be based on legal, biopharmaceutic, pharmacokinetic, pharmacologic and clinical considerations.

From a legal standpoint, the US Food and Drug & Cosmetic Act states:

> A drug shall be considered to be bioequivalent to a listed drug if (i) the rate and extent of absorption of the drug do not show a significant difference from the rate and extent of absorption of the listed drug.
>
> FD&C Act 505(j) (7) (B)

From a biopharmaceutic standpoint, discharge and dissolution of an active drug substance from a pharmaceutical formulation comprise a series of complex events occurring in the gastrointestinal tract that are not easily reflected in a single parameter, such as k_a. Similarly, from a pharmacokinetic standpoint, the definition of a rate constant of absorption, such as k_a, from the small number of pre-peak data points in a plasma concentration—time curve is frequently difficult. From a clinical pharmacology standpoint, comparison of the hypothetical rate constants might be less important than determination of comparability in pharmacologic effect.

Consider the following statements:

Drug Information Association statement (9)

> Two pharmaceutical products are considered to be equivalent when their bioavailabilities from the same molar doses are so similar that they are unlikely to produce clinically relevant differences in therapeutic and/or adverse effects.

Statement from Sheiner (10)

> The main point is that the logical basis for current bioequivalence measurement and regulation is seriously inadequate; only with an appropriate model for dose effect, and a clear delineation of clinical context and values, can one devise, estimate and test bioequivalence measures that make clinical and scientific sense.

These positions reflect to some degree a continuum of thought, starting with a legal definition, moving through var-

ious scientific disciplines (biopharmaceutics and pharmacokinetics), and ending with what must be the most important objective—clinical efficacy and safety comparisons that allow therapeutic interchangeability. The challenge to the regulatory scientists is to formulate a rational public policy to assess formulation bioequivalence that addresses these different perspectives.

Thee is apparently no universal measure of rate. All rate measures studied fail in one aspect or another. We certainly cannot generally recommend the use of C_{max}/t_{max}, C_{max}/AUC_{tmax}, SL_f or AUC_f. They have very high variability and give irreproducible results across clinical trials.

The measure t_{max} has a good sensitivity, but yields very high producer risks. The use of a smoothed or model-estimated t_{max} might ameliorate its behavior. Its power curve could also be improved by using statistical procedures other than the t-test (e.g., nonparametric tests, more adapted to the discrete nature of t_{max}). However, no such tests have been validated for bioequivalence use.

The C_{max} measure has low producer risk, but also low sensitivity. Between two formulations, differences in rate well above 25% often go unnoticed when using C_{max}. This sets a *de facto* lower standard on rate, since C_{max} is commonly used. A 25% difference in C_{max} corresponds to a larger percentage difference in rate (k_a), unless extent (to which C_{max} is sensitive) is also different between the test and the reference formulation.

The C_{max}/AUC_∞ ratio is an attempt to make C_{max} independent of extent (which could simplify decision making). When the fraction absorbed, F, is highly variable, this measure should perform better than C_{max}. When F is low C_{max} may be a better measure, especially in the case of "flip-flop" kinetics and when two-compartment kinetics induce a poor behavior of the AUC component of the ratio. In the latter situation measures such as C_{max}/AUC_{tqc} may have better behavior (1). It is however unlikely that sensitivity will improve dramatically.

The partial area up to t_{max} (AUC_p) for the reference formulation, has very satisfying behavior (low producer risk, higher sensitivity than C_{max} or C_{max}/AUC_∞, except for a random lag-time, where it fails to ever show bioequivalence. This behavior might be improved by modifying its operational definition. However, care should be taken to keep it sensitive to nonrandom lag-times (i.e., higher average lag for one formulation relative to the other), since such a lag leads to a difference in rate.

The use of multiple concentrations themselves, or cumulative AUCs, should be further explored. The different time points have different sensitivities and variabilities, depending on the scenario. To avoid arbitrary choices of time, a whole set of time points should be used, but adequate statistical procedures are still lacking do so. In addition, the performance of a new measure (11), assessing the "distance" between test and reference time-concentration curves, should also be evaluated.

According to the US statutory requirement "rate" and "extent" of absorption must be documented. We have assumed that a 25% difference in rate or extent of absorption between two different formulations is of interest (1). While this interest is probably reasonable for extent, it is immediately apparent, based on our simulations, that it is not reasonable for rate. Small differences in absorption rate on the order of 25% are not likely to impact greatly on the concentration-time profile for most drugs, which the pharmacokineticists might argue is our primary concern. From the perspective of the pharmacokineticist, the question could be the reverse: if formulations are not comparable and produce, for example, a 25% differences in peak concentration, what magnitude of rate change must occur to yield such a difference?

While pharmacokineticists might argue that a parameter determining comparability in "rate" of absorption is not pertinent, the clinicians and clinical pharmacologists might equally argue that comparability in measures of a plasma concentration-time curve are also not pertinent—that the primary comparison of interest should be comparability in pharmacologic effects between two formulations. This argument carries the determination of bioequivalence to a different level, where information about dose-response and concentration-response relationships for any of several efficacy and toxicity effects of a drug become important. This information might then be used to set equivalence criteria depending on the dose-response information available for a given drug. Unfortunately, information about these relationships is absent for many, if not most, drugs. Even when it is available, the data almost always show substantial variability compared to the concentration of a drug or metabolite in an accessible biological fluid.

The justification for relying on comparability in plasma (or urine) concentration-time curves, where available, becomes a reasonable and reliable method of documenting bioequivalence for most drugs. If one accepts this focus, namely, that plasma concentration-time curve data should drive most determinations of bioequivalence, comparability in these levels becomes paramount, as opposed to comparisons of pharmacodynamic parameters or a hypothetical rate constant. Where available and as necessary, pharmacodynamic data might be used to widen equivalence criteria for Cmax or to require comparison of concentrations or areas during the initial portion of a concentration-time curve. Absent this information, regulatory requirements would default to the conservative approach that states that formulations are comparable and therapeutically interchangeable if differences in measures from a plasma concentration-time curve, such as C_{max} and AUC, do not exceed 25%. In this report, we also document that other measures beyond C_{max} might be more appropriate to assess the "rate" component of the concentration-time curve, depending on the underlying pharmacokinetic characteristics of the drug. Irrespective of which measure is chosen to reflect the input rate of the drug, the primary focus becomes one of assuring that the patterns of the plasma concentration-time profiles are comparable between the two formulations, on the premises that comparability in these patterns will lead to therapeutic equivalence.

CONCLUSION

The inability of the various measures tested to detect a 25% change in rate can be viewed from two different perspectives. First, if k_a is used as the criterion, then one would need to determine the change in a given measure corre-

sponding to a specified change in k_a. However, this is problematic as the criterion for the measure depends on the formulation and kinetic properties of the drug. Current practice corresponds to the reverse; by fixing a criterion for C_{max}, the implied criterion for k_a is wider than that for C_{max} by an amount depending on the formulation and the drug. Second, from a pharmacokinetic point of view, controlling for k_a (or the rate profile) itself is useful when the therapeutic indication requires a rapid onset of action. The problem is that no currently available rate measure is sensitive enough to detect a 25% difference in k_a. On the other hand, for most drugs rate is not so important; extent and peak plasma levels are the major concerns. Furthermore, the rate-time profile, even for a given formulation, is often so variable that requiring only a 25% difference in the absorption rate constant may be unrealistic. Thus, controlling C_{max} seems appropriate.

These studies have shown that there is no universal measure for rate of absorption. They have also demonstrated that each rate measure has advantages and limitations that depend on the kinetic properties of the drug and its formulation. Consequently, simulations of bioequivalence testing are highly recommended to assess the applicability of the rate measures within the context of a specific situation.

ACKNOWLEDGMENTS

This work was supported by inter-agency agreement #224-91-3007 between the US Food and Drug Administration and the US Department of Energy.

REFERENCES

1. 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.* (in press).
2. W. W. Hauck and S. Anderson. Types of bioequivalence and related statistical considerations. *Int. J. Clin. Pharmacol. Ther. Toxicol.* 30:181–187 (1992).
3. M. L. Chen. An alternative approach for assessment of rate of absorption in bioequivalence studies. *Pharm. Res.* 9:1380–1385 (1992).
4. L. Endrenyi, S. Fritsch, and W. Yan. Cmax/AUC is a clearer measure than Cmax for absorption rates in investigations of bioequivalence. *Int. J. Clin. Pharmacol. Ther. Toxicol.* 29:394–399 (1991).
5. M. Rowland and T. N. Tozer. *Clinical Pharmacokinetics: Concepts and Applications,* Lea & Febiger, Philadelphia, 1989.
6. A. Racine-Poon and A. F. Smith. Population models. In D. A. Berry (ed.), *Statistical Methodology in the Pharmaceutical Sciences,* Marcel Dekker, Inc., New York, 1990, pp. 139–162.
7. D. J. Schuirmann. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J. Pharmacokin. Biopharm.* 15:657–680 (1987).
8. Food and Drug Administration—Office of Generic Drugs—Division of Bioequivalence. Guidance on statistical procedures for bioequivalence studies using a standard two-treatment crossover design. 1992.
9. Drug Information Association Working Group. International harmonization and consensus DIA meeting on bioavailability and bioequivalence requirements and standards. *Drug Info. J.* 25:471–482 (1991).
10. L. B. Sheiner. Bioequivalence revisited. *Statistics in Medicine* 11:1777–1788 (1992).
11. A. Rescigno. Bioequivalence. *Pharmaceutical Research* 9:925–928 (1992).