Why Rate of Absorption Inferences in Single Dose Bioequivalence Studies are Often Inappropriate

Rodney P. Basson,[1,2] Atalanta Ghosh,[1,3] Benito J. Cerimele,[1] Karl A. DeSante,[1] and Daniel C. Howey[1]

Received August 25, 1997; accepted October 31, 1997

Purpose. Peak drug concentration (C_{max}) measures the extremity of drug exposure and is a secondary indicator of the extent of absorption after area under the concentration time curve (AUC). C_{max} serves as *the* indicator of absorption *rate* in bioequivalence (BE) studies in the US (1). The use of C_{max}, not the time to C_{max}(T_{max}), as the metric to assess absorption rate causes erratic inferences in BE studies, and incorrect conclusions for some. We can improve BE efficiency (i.e., get the answer right the first time), by properly analyzing the time to C_{max}(T_{max}) instead of C_{max}.

Methods. We have previously redirected attention to T_{max} as *the* unconfounded absorption rate variable, instead of C_{max}, and have called for equally spaced sampling times during the suspected absorption phase to improve the performance of the rate metric (2). Equal spacing converts T_{max} easily into a count variable and we illustrated an appropriate statistical analysis for counts. This paper provides some measurement theory concepts to help judge which is the more appropriate analysis, and also provides parametric confidence limits for T_{max} treatment differences. Three separate BE studies are then analyzed by both methods.

Results. By focusing on the differences in conclusions, or inferences, this paper identifies three major issues with the current FDA "recommended" analysis of BE studies. First, C_{max}, a continuous variable peak-height or extent measure has usurped T_{max}'s function and performs erratically as a substitute measure for the rate of absorption. Second, T_{max}, should be analyzed as a discrete attribute, not as a continuous variable. Third, since several extent measures (AUC, C_{max}), not one, are actually being analyzed, an adjustment for multiple testing is mandatory if we are to maintain the size of the test at the desired α level (13), and not inadvertently use a narrower bioequivalence window than is intended. These actions all *can* have serious unintended consequences on inferences, including making inappropriate ones.

KEY WORDS: bioequivalence; absorption rate; Tmax; absorption process rate; measurement theory; inference.

INTRODUCTION

Generic competition in human drugs is a laudable public interest goal with which, in principle, few disagree. A finding of bioequivalence (BE) serves as a surrogate for therapeutic identity (1), so a BE evaluation (BEE) process that is fair and efficient should foster generic drug competition. Because companies use BEE to also assess formulation changes during drug development, an efficient process grounded in statistically sould science is important. BE is customarily evaluated *in vivo* in healthy subjects, by comparing both *the rate and the extent* of drug absorption of a test formulation with a reference formulation. The area under the concentration time curve from time zero to time t (AUC_{0-t}, where t is the last measurable time point) and similarly the area under the curve from time zero to time infinity ($AUC_{0-\infty}$) are both accepted as uncontaminated measures of the extent of absorption. The situation for rate is in a state of confusion. There are drugs whose plasma concentrations are plateaued, where the concept of rate of absorption is not well defined, and ought therefore not be calculated. In regular time–concentration profiles, data on time to peak absorption, T_{max}, are typically collected but with irregular sampling schemes. Such data beg the question, and are not routinely amenable to proper statistical evaluation (3). The continuous variable C_{max}, the highest observed concentration, and undeniably also a measure of extent, has quietly usurped the function of T_{max} and performs as a surrogate measure for the rate of absorption. A formulation problem that C_{max} is uniquely qualified to evaluate is 'dose-dumping', i.e. reaching an unsafe concentration, and retaining it to evaluate safety peripherally for that purpose, is not in dispute. When in addition to C_{max}, T_{max} is analyzed, often it too is regarded as a continuous variable. There are consequences to all these actions, and we illustrate how erratic, error prone rate inference can afflict BEE by way of several pertinent examples.

METHODS

Absorption Phase Sampling Times

To assess the rate of absorption we advocate the use of *equal* spacing of the sampling times from time zero (or other suitable initial time) until approximately two or three times the expected peak concentration time (or other suitable *absorption* phase restricted time interval) to collect pertinent rate data. For example, a drug which has a T_{max} of approximately 30 minutes in fasted subjects, and a similar short elimination half-life, is easily densely sampled every fifteen minutes for the first two hours (nine samples) and with diminishing frequency thereafter through eight hours. An equal sampling interval through the absorption phase ensures that a subject's T_{max} in hours multiplied by the sampling frequency per hour is always a positive integer. These integer counts tell how long the absorption process takes to reach maximum concentration for each subject, so they encapsulate the process rate (2).

Measurement Theory

Measurement Theory (MT) is a branch of applied mathematics with serious implications for data analysis (4). MT instructs that measurements (i.e., the data) and the attributes being measured (i.e., the reality represented like absorption rate) are not one and the same; to draw valid conclusions about an attribute one must take into account the nature of the correspondence between the attribute and its measurement. In particular, MT says if statistical inferences made from measurements are to apply to reality, then we must pay attention also to other key MT tenets such as the 'level of measurement' and

[1] Lilly Laboratory for Clinical Research, Eli Lilly and Company, Indianapolis, Indiana 46285.

[2] To whom correspondence should be addressed.

[3] To whom a request for the confidence interval macro program should be addressed.

0724-8741/98/0200-0276$15.00/0

'permissible transformations'. For example, because integer counts reflect an absolute level of measurement, T_{max} data that arise for subjects from a crossover study need to be analyzed appropriately. Measurement theorists warn the price usually paid for inattention to these matters is 'meaningless statements'. Uncritical analysis of C_{max} and T_{max} data for subjects in a BEE, is a practical example of such inattention.

Statistical Methodology

An analysis of T_{max} count data for subjects from a crossover study and which meets these MT constraints has been published (2). Performing the computations for counts that arise within the context of a generalized linear model is an innovation largely due to McCullagh and Nelder (5). Software that can perform the necessary computations is available from SAS Institute Inc. (6). Software to perform exact nonparametric inference for count data is available from Cytel Software Corp. (7). Using the appropriate variance–covariance matrix of the estimates to calculate confidence limits for estimates, and their differences, the analysis given in (2) can be extended. This report illustrates this for three BE studies.

Continuous data for subjects that arise from a crossover study, like C_{max} and AUC, have long been analyzed by general linear model procedures. Westlake (8) first suggested the log transforms of C_{max} and AUC be analyzed, and a confidence interval method similar to that given by Schuirman (9) is used to compare values between two formulations. A comprehensive approach to performing these computations is described in SAS Institute Inc. (10,11). Schultz and Steinijans (12) have advocated widening the bioequivalence range for C_{max} from 0.8–1.25 to 0.7–1.43, and this has been embraced by Canadian and EEC authorities. Nonparametric bounds for T_{max} which are similar to those we present, have been published (12).

Alternative approaches are usually based on C_{max}. While traditional, the practice of analyzing C_{max} data instead of T_{max} for absorption rate simply is not logical (2). Widening the BE bounds for C_{max}, but not for AUC, is a empirical solution, and preferable to doing nothing. It seems not to be generally perceived that multiple testing is involved. For example, if AUC_{0-t} is chosen as the primary extent variable then it follows that $AUC_{0-\infty}$ is a secondary extent C_{max} is a tertiary extent variable.

If all these extent variables are to be analyzed in a BE study, then some form of adjustment for multiple testing is mandatory if we are to maintain the size of the test at the desired α level (13); otherwise a narrower bioequivalence window than was intended is used.

Three Ordinary BE Studies

Single dose crossover studies (two antibiotic, one antiviral) were conducted by Lilly Research Laboratories to assess BE of test and reference formulations. Periods of blood collection lasted at least eight hours, after a single dose, and treatment periods were separated by a washout period of not less than three days between doses. Twenty four healthy male subjects enrolled in and completed both study 1 and study 3. Sixteen healthy male subjects enrolled in study 2; selected pharmacokinetic data appeared in (2).

Analyses performed in accordance with standard BE guidelines was applied to logarithmic transformations of three variables: C_{max}, AUC_{0-t} and $AUC_{0-\infty}$ respectively.

RESULTS

Comparison of Analysis Conclusions for Absorption Rate

Conventional BEEs for the three studies appears in Table I. In all three studies AUC declares formulations to be BE in the extent of absorption. As is so often the case in BE, if there is a 'bad actor' it will be C_{max}. However, in all three examples, C_{max} passes the widened (.70 to 1.43) BE bounds advocated for C_{max} by Canadian and EEC authorities, good evidence here that none of these studies really have a bioinequivalence problem.

In study 1 two additional variables, T_{max} and half-life, were analyzed by questionable continuous variable significance tests. The tests all agreed that the new formulations and the marketed product were BE. When T_{max} means are very close, as in this example, they are unlikely to be found statistically different in absorption rate, no matter what assessment

Table I. 90% Confidence Limits on Separation of Least Squares Mean (Log transformed Data)

Contrasts[a]	Least Squares Mean: Test	Least Squares Mean: Reference	Ratio of means[b]	90% Confidence Interval[c]	Outcome
Study 1-Antibiotic #1					
A vs. C					
C_{max}	17.25	17.57	.97	.86 to 1.11	Pass
AUC_{0-t}	29.55	29.37	1.00	.94 to 1.06	Pass
$AUC_{0-\infty}$	29.84	29.64	1.00	.94 to 1.06	Pass
B vs. C					
C_{max}	18.72	17.57	1.06	.93 to 1.21	Pass
AUC_{0-t}	29.36	29.37	1.00	.95 to 1.06	Pass
$AUC_{0-\infty}$	29.63	29.64	1.00	.95 to 1.06	Pass
Study 2-Antibiotic #2					
A vs. C					
C_{max}	13.97	14.94	0.95	0.82 to 1.12	Pass
AUC_{0-t}	13.18	13.30	0.99	0.93 to 1.04	Pass
$AUC_{0-\infty}$	13.43	13.54	0.99	0.94 to 1.04	Pass
B vs. C					
C_{max}	12.65	14.94	0.82	0.70 to 0.97	Fail
AUC_{0-t}	13.53	13.30	1.01	0.95 to 1.07	Pass
$AUC_{0-\infty}$	13.78	13.54	1.01	0.96 to 1.07	Pass
Study 3-Antiviral					
A vs. B					
C_{max}	321.3	362.9	.89	0.75 to 1.05	Fail
AUC_{0-t}	1488.3	1531.3	0.97	0.85 to 1.10	Pass
$AUC_{0-\infty}$	1657.5	1637.3	1.01	0.93 to 1.10	Pass

[a] Units for parameters: C_{max}, ng/mL; AUC_{0-t} and $AUC_{0-\infty}$, ng·hr/mL.
[b] Analyses of C_{max} and AUC parameters are based on log-transformed data. Antilogs of transformed scale fed minus fasted differences and their 90% confidence limits supply a test/reference ratio estimate and corresponding 90% confidence interval. The point estimate of the ratio of equivalent means is 1.0.
[c] BE range is 0.80 to 1.25.

method we use, and this study illustrates the fact that C_{max} can find the correct answer by sheer luck.

In study 2 the C_{max} result appears due to three subjects in particular. On formulation B all three had unexpectedly *low* C_{max} readings, 7.3 μg/ml, 7.8 μg/ml and 7.1 μg/ml, respectively, and correspondingly long T_{max} values, 1.25 hr, 1.75 hr and 1.25 hr, respectively, for a drug that usually is absorbed in 1 hr or less (2). A plausible explanation for this anomaly is that the fasting state was compromised.

Analysis of T_{max} regarded as a continuous variable in this study actually contradicts C_{max} and found no statistically significant differences ($p = 0.13$) beyond random fluctuation between the B and C formulations. However, this indication would usually be overruled, and the study would be repeated.

In study 3 the C_{max} result is caused by subject 10 whose low concentration time profile on formulation A was noticeably aberrant. Corresponding T_{max} values for subject 10 were not affected, so T_{max} treated as a continuous variable is robust to this influential subject (p-value = 0.61). No matter, this study also, would be repeated.

The discrete T_{max} count data for the three studies is given in Table II; an appropriate analysis for this absolute variable appears in Table III.

In studies 1 and 2, which both used a sample collection interval of 15 minutes, absorption rate means have an estimated standard deviation of 18 minutes and 19 minutes, respectively. The mean formulation absorption rates in study 3 have a standard deviation of 27 minutes. This too is slightly greater than the sampling density interval (1/3 hour, or 20 minutes) used to collect the data.

The reciprocal of the constant absorption phase sampling rate seems to roughly define a standard set of confidence bounds for absorption rate estimates. When estimates differ by less than this basic indeterminacy interval they are statistically indistinguishable from each other. Analysis of C_{max} is necessarily oblivious to this fundamental limitation imposed by the study design. So in contrast to the conventional analysis, in all three cases, Poisson regression analysis of T_{max} fails to declare any of the *formulations* to have different absorption rates. This more robust analysis happens to agree with the 'widened bounds' approach, in these three examples.

DISCUSSION

These three BE examples are not extraordinary. They illustrate that tightly regulated BE is not a fair game of chance. In the first example C_{max} is barely well-behaved enough to not disrupt an otherwise uneventful study. In study 2 a more common BE dilemma occurs. A few subjects who were supposed to observe a 14 hr fast exhibit signs consistent with having consumed food. The rate surrogate C_{max} admirably detects data outliers, but, if used inappropriately, will declare that *formulations* B and C are not BE. Current regulatory guides are inflexible and intolerant regarding biological variation or individual outliers so that all an unfortunate sponsor can do is bury the study, and repeat the work hoping for better luck. Analyses which deal more realistically with individual subject variability, declare B and C formulations BE in both rate and extent of absorption! In the second and third studies even naive analysis of T_{max} actually does contradict an errant C_{max}; apparently however that signifies nothing.

There are latent unintended consequences to the BE

Table II. Number of Fractional Hours to Reach Tmax

Study 1-Antibiotic #1

	Frequency of T_{max} value by sampling time (quarter hr.)								
Formulation	1	2	3	4	5	6	7	8	Total
A	0	1	8	5	9	1	0		24
B	0	2	10	4	4	2	2		24
C	0	0	9	6	5	4	0		24
Total	0	3	27	15	18	7	2		72

Study 2-Antibiotic #2

	Frequency of T_{max} value by sampling time (quarter hr)								
Formulation	1	2	3	4	5	6	7	8	Total
A	2	9	4	0	0	0	0	0	15
B	2	6	4	0	2	0	1	0	15
C	3	8	3	1	0	0	0	0	15
Total	7	23	11	1	2	0	1	0	45

Study 3-Antiviral

	Frequency of T_{max} value by sampling time (third hr)								
Formulation	1	2	3	4	5	6	8	9	Total
A	0	2	5	4	6	3	2	1	24
B	0	1	6	4	6	5	3	0	24
Total	0	3	11	8	12	8	5	1	48

Table III. Histograms Means and Confidence Intervals (Poisson Regression)

Contrastsa	Mean Test	Mean Reference	Difference in means	90% Confidence Interval	Significance p-value
Study 1-Antibiotic #1					
A vs. C					
T_{max}	60	62	-2	-17 to 13	0.83
B vs. C					
T_{max}	61	62	-1	-16 to 13	0.89
Study 2-Antibiotic #2					
A vs. C					
T_{max}	32	32	0.0	-14 to 14	1.0
B vs. C					
T_{max}	43	32	11	-4 to 26	0.20
Study 3-Antiviral					
A vs. B					
T_{max}	1.54	1.63	-0.08	-0.44 to 0.27	0.69

[a] Units for parameters: T_{max}, min. for studies 1 and 2, hr for study 3.
[b] The Test minus reference difference is given. The point estimate of the difference in equivalent means 0.0.

dilemma. To control the rogue C_{max} variable, sponsors run ever larger, costlier, studies. Subjecting two dozen or more subjects to invasive experimentation has become routine in BEE. Naive analysis of T_{max}, particularly in a large study, generates its own problems.

CONCLUSIONS

Efficient BEE grounded in good science will facilitate fair competition on a flat playing field and is a worthy goal.

We have illustrated a way to improve the overall efficiency of the BEE process. The first step consists in using a refined but simple sampling scheme to empower T_{max} to function as the metric for absorption rate. The second step requires no more than respect for the tenets of measurement theory: analyze the absorption rate attribute, T_{max}, (a count variable), with an appropriately restricted statistical analysis.

Alternative approaches are more convoluted. If AUC is the primary extent measure, then admit C_{max} is a *secondary* extent variable. To analyze C_{max} and control the α level of the test, make an adjustment for the multiple testing that this action implies (13). Widening of the BE bounds arbitrarily for C_{max}, but not for AUC, is an alternative approach that has already been embraced by Canadian and EEC authorities.

This will lower the incidence of inappropriate inferences in BE studies, reduce the need to redo many of them, save that development cost, and reduce the number of subjects studied in each BEE.

REFERENCES

1. Division of Bioequivalence Guidance Statement. *Statistical procedures for bioequivalence studies using a standard two-treatment crossover design.* Office of Generic Drugs, Food and Drug Administration, Rockville MD, 1992, pp. 1–13.
2. R. P. Basson, B. J. Cerimele, K. A. DeSante, and D. C. Howey. *Pharmaceutical Research,* **13**:324–328. 1996.
3. W. Feller. *An Introduction to Probability Theory and its Applications. 2nd edition.* John Wiley and Sons, New York, 1957, pp. 135–163.
4. W. S. Sarle. *Measurement Theory: Frequently asked questions.* Disseminations of the International Statistical Applications Institute, 4th Edition, ACG Press, Inc., Wichita KS, **1**:pp. 61–66. 1995.
5. P. McCullagh and J. A. Nelder. *Generalized Linear Models.* Chapman and Hall, London, 1989.
6. SAS Institute: *SAS/STAT™ Software: The Genmod Procedure, Release 6.09.* SAS Institute, Cary, NC, pp. 1–88. 1993.
7. Cytel Software Cor.: *StatXact. Statistical Software for Exact Nonparametric Inference,* Cytel Software Corp., Cambridge, MA 02139, Chapter 4. 1991.
8. W. J. Westlake Bioavailability and bioequivalence of pharmaceutical formulations. In: K. E. Peace (eds.), *Biopharmaceutical Statistics for Drug Development,* Marcel Dekker, New York, pp. 329–352. 1988.
9. D. J. Schuirmann. *J. Pharmacokin. Biopharm.* **15**:657–680. 1987.
10. SAS Institute: *SAS/STAT™* User's Guide, Release 6.3 Edition. SAS Institute, Cary, NC, 1988, pp. 549–640.
11. SAS Institute: *SAS™ Technical Report, SAS/STAT™ Software: Changes and Enhancements, Release 6.07,* SAS Institute, Cary, NC, pp. 371–405. 1992.
12. H.-U. Schultz and V. W. Steinijans. *Int. J. Clin. Pharmac. Ther. Toxic.* 30 (Suppl. 1), 1992, S1–S6.
13. P. H. Westphal and S. S. Young. *Resampling-Based Multiple Testing: Examples and Methods for p-value Adjustment.* Wiley-Interscience, New York, NY. 1993.