Korean J Physiol Pharmacol
Vol 13: 367−371, October, 2009
DOI: 10.4196/kjpp.2009.13.5.367

Comparison of Parametric and Bootstrap Method in Bioequivalence Test

Byung-Jin Ahn, and Dong-Seok Yim

Department of Pharmacology, College of Medicine, The Catholic University of Korea, Seoul 137-701, Korea

The estimation of 90% parametric confidence intervals (CIs) of mean AUC and Cmax ratios in bioequivalence (BE) tests are based upon the assumption that formulation effects in log-transformed data are normally distributed. To compare the parametric CIs with those obtained from nonparametric methods we performed repeated estimation of bootstrap-resampled datasets. The AUC and Cmax values from 3 archived datasets were used. BE tests on 1,000 resampled datasets from each archived dataset were performed using SAS (Enterprise Guide Ver.3). Bootstrap nonparametric 90% CIs of formulation effects were then compared with the parametric 90% CIs of the original datasets. The 90% CIs of formulation effects estimated from the 3 archived datasets were slightly different from nonparametric 90% CIs obtained from BE tests on resampled datasets. Histograms and density curves of formulation effects obtained from resampled datasets were similar to those of normal distribution. However, in 2 of 3 resampled log (AUC) datasets, the estimates of formulation effects did not follow the Gaussian distribution. Bias-corrected and accelerated (BCa) CIs, one of the nonparametric CIs of formulation effects, shifted outside the parametric 90% CIs of the archived datasets in these 2 non-normally distributed resampled log (AUC) datasets. Currently, the 80∼125% rule based upon the parametric 90% CIs is widely accepted under the assumption of normally distributed formulation effects in log-transformed data. However, nonparametric CIs may be a better choice when data do not follow this assumption.

Key Words: Bioequivalence, Bootstrap, Confidence interval, Nonparametric method

INTRODUCTION

Average bioequivalence (ABE) is commonly tested for PK parameters (e.g. AUC and Cmax) obtained from BE studies of crossover design. Generally, log (AUC) and log (Cmax) values are statistically analyzed using the mixed effect or two-stage linear model. Based on the two-one sided test (TOST) proposed by Shuirmann (1987) (FDA, 2001), two formulations are claimed to be bioequivalent when the 90% confidence intervals (CIs) of mean log (AUC) differences and log (Cmax) differences fall within the regulatory acceptance limits (log (0.8) to log (1.25)). The mean differences in log (AUC) or log (Cmax) between the test and the reference formulation represent the formulation effect, a key parameter in the ABE test (Patterson and Jones, 2002).

Because AUC and Cmax are positive and right-skewed, they have been regarded as log normally-distributed (Midha et al., 1993; Chow, 2003). Nonparametric methods may be indicated for data which do not follow a normal distribution even after some transformation. However, because of the poor sensitivity of nonparametric procedures for small data, other more reliable methods are needed (Pabst and Jaeger, 1990). Since it is unknown what the distribution of formulation effects for AUC and Cmax is like, we decided to investigate the distribution pattern of formulation effect estimates of 1,000 bootstrap-resampled datasets from each BE study dataset. In other words, after the formulation effect estimate for each resampled dataset was obtained by BE tests using SAS, the distribution pattern of 1,000 such estimates was analyzed instead of assuming it to be log-normal. Although the resampling approach is an approximation accepting an assumption that may not be true, it is a useful alternative (Henderson, 2005). Then, the 90% nonparametric CIs for log (AUC) and log (Cmax) differences between 2 formulations in the 1,000 resampled datasets were estimated using several different non-parametric methods. The nonparametric CIs were then compared with the 90% CIs obtained from BE tests (parametric CIs) on the archived datasets.

METHODS

Datasets of AUC and Cmax from 3 BE studies (named as BE1, BE2 and BE3 henceforth) previously conducted in the institution were used for this study. The numbers of subjects for BE1, BE2 and BE3 were 23, 23 and 24,

Received August 19, 2009, Revised September 1, 2009, Accepted September 22, 2009

Corresponding to: Dong-Seok Yim, Department of Pharmacology, College of Medicine, The Catholic University of Korea, 505, Banpo-dong, Seocho-gu, Seoul 137-701, Korea. (Tel) 82-2-2258-7888, (Fax) 82-2-2258-7859, (E-mail) yimds@catholic.ac.kr

ABBREVIATIONS: AUC, area under the time-plasma concentration curve; Cmax, peak plasma concentration; Tmax, time to reach the Cmax; BE, bioequivalence; CI, confidence interval; BC, bias-corrected; BCa, bias-corrected and accelerated.

respectively. All 3 studies used the standard two-sequence, two-period crossover design. Results of BE1 and BE3 satisfied the BE criteria, but that of BE2 did not.

Estimation of formulation effects and their standard errors in the 3 aforementioned archived datasets was performed using SAS (Enterprise Guide version 3.0, SAS Institute Inc, Cary, North Carolina, USA). Bootstrap-resampling was repeated 1,000 times using the MACRO function of SAS for each archived dataset. BE tests were also performed for these 1,000 resampled datasets to find the distribution of the point estimates of formulation effects. For this purpose, normal Q-Q (Quantile to Quantile) plots, histograms and density curves were plotted, and skewness and kurtosis therein were also calculated. Shapiro-Wilk tests were also performed to obtain p values of the normality of the distribution. Several nonparametric 90% CIs of formulation effects in bootstrap-resampled datasets were then compared with the parametric 90% CIs obtained from BE tests on the 3 archived datasets. We estimated nonparametric CIs including percentile CI, bootstrap-t CI, Bias-corrected (BC) CI and Bias-corrected and accelerated (BCa) CI (Efron and Tibshirani, 1993; Bonate, 2005).

To compare the lengths of nonparametric and parametric 90% CIs, we adopted the term "percent coverage" which was mentioned by Bonate (Bonate, 2005). The percent coverage is 90% and 45% when the length of the nonparametric CI is identical to or half of the parametric CI, respectively.

RESULTS

Fig. 1 illustrates the distributions of estimated formulation effects of log (AUC) in bootstrap-resampled datasets using histogram-density curves and normal Q-Q (Quantile to

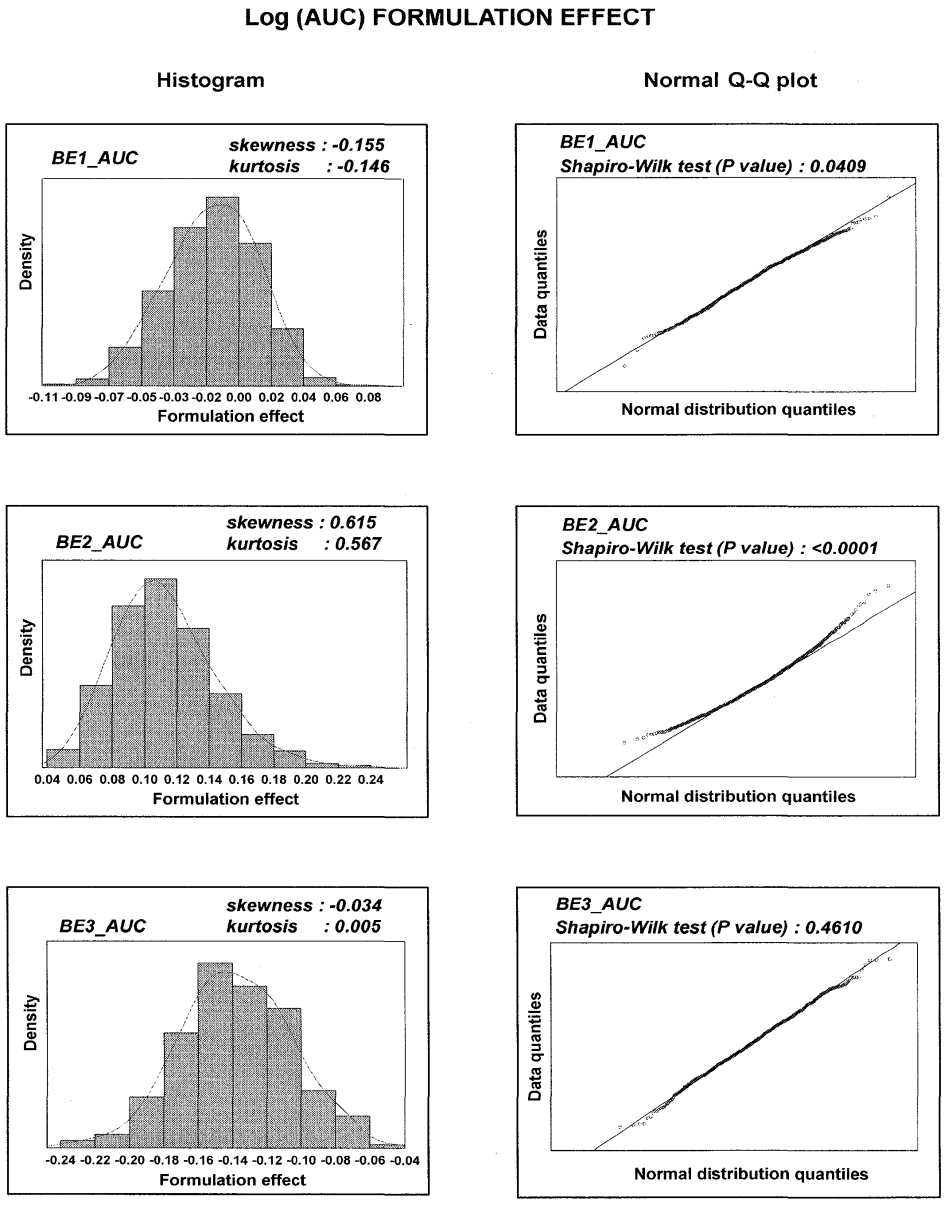

Fig. 1. Histograms and normal Q-Q plots with skewness, kurtosis and Shapiro-Wilk test results showing the distribution of formulation effects (mean differences in log (AUC)s between test and reference formulation) estimates in bootstrap-resampled datasets.

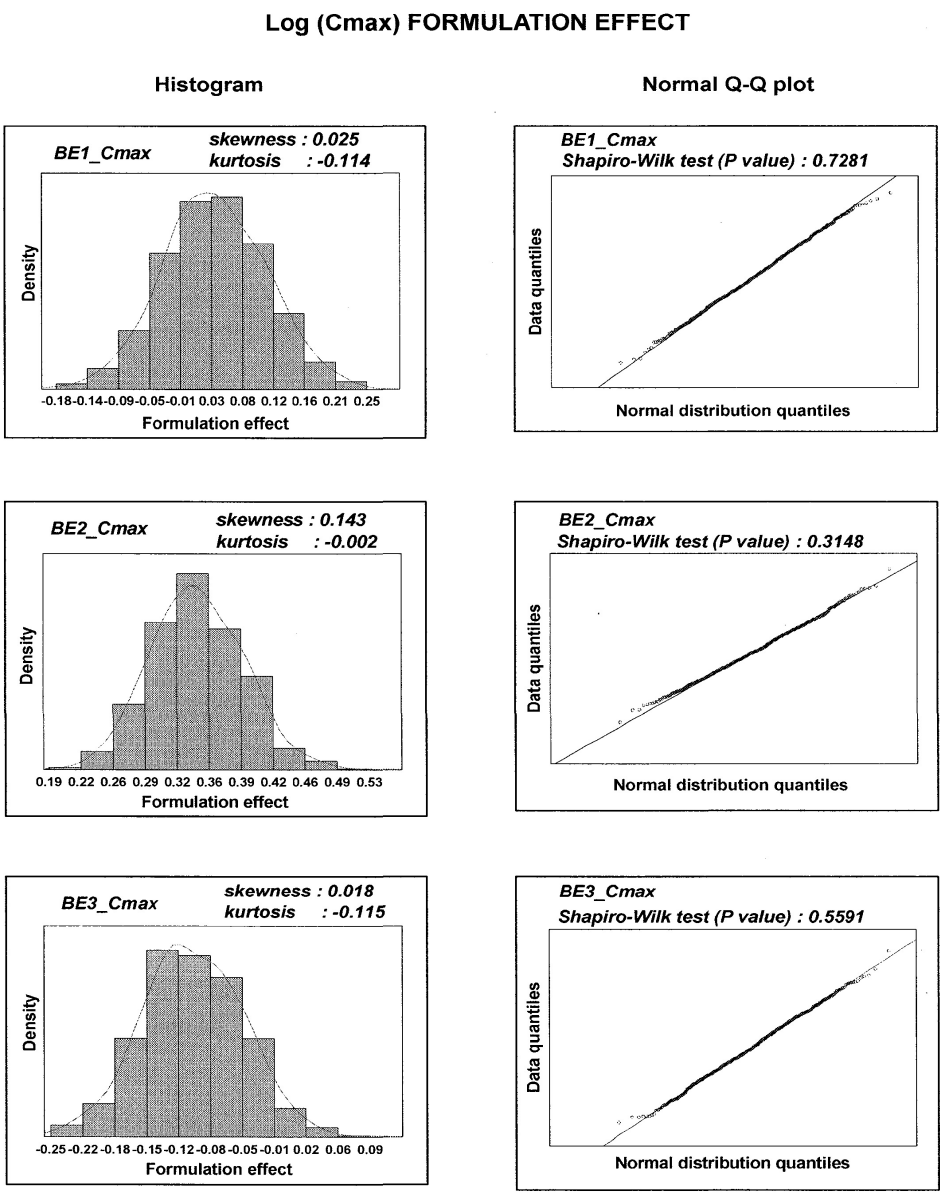

Fig. 2. Histograms and normal Q-Q plots with skewness, kurtosis and Shapiro-Wilk test results showing the distribution of formulation effects (mean differences in log (Cmax) between test and reference formulation) estimates in bootstrap-resampled datasets.

Quantile) plots. Shapiro-Wilk normality test results are also given in the panels. The formulation effects for BE2_AUC showed a right-skewed distribution. Although the skewness of the formulation effect distribution for BE1_AUC was not visually evident, it did not pass the Shapiro-Wilk test ($p < 0.05$) (Fig. 1). In the case of Cmax, all histograms for the formulation effect distribution passed the Shapiro-Wilk test (Fig. 2). When we compared the 4 different nonparametric bootstrap 90% CIs of formulation effects with the parametric 90% CIs from the archived datasets, there were slight differences between them. The nonparametric percentile 90% CIs showed lower percent coverages for all 3 BE studies (80～86%). The bootstrap-t 90% CIs were wider than parametric CIs for BE1 and BE2 studies (about 93～112%) but not for BE3 (about 87%). Bootstrap BC and BCa 90% CIs had percent coverage less than 90% (79～88%) for all 3 BE studies (Table 1).

In 2 cases where the distribution of formulation effects did not pass the Shapiro-Wilk normality tests (BE1_AUC, BE2_AUC), their BCa 90% CIs shifted over the boundaries of the parametric 90% CIs (BE1_AUC andBE2_AUC). However, the other 4 cases with normally distributed formulation effects showed slightly narrower ranges without the shifts observed in BE1_AUC and BE2_AUC (Fig. 3).

DISCUSSION

The basic assumption of the parametric BE test is that log-transformed AUC and Cmax are normally distributed. If this assumption is not true, BE may have to be tested non-parametrically. In this study, we attempted to test the log-normality of the formulation effect using bootstrap-resampling methods. There are several ways to determine

Table 1. The 90% confidence intervals (CI) of formulation effects in the archived and bootstrap-resampled datasets and percent coverages of nonparametric CIs in contrast with parametric CIs

		Log (AUC)					Log (Cmax)				
		Parametric	Nonparametric				Parametric	Nonparametric			
Dataset	Parameter	Archived-t	Percentile	Bootstrap-t	BC	BCa	Archived-t	Percentile	Bootstrap-t	BC	BCa
BE1	5%	−0.059	−0.056	−0.064	−0.059	−0.06	−0.083	−0.077	−0.097	−0.07	−0.071
	95%	0.033	0.028	0.031	0.026	0.027	0.162	0.157	0.171	0.161	0.162
	90% Interval	0.092	0.084	0.095	0.085	0.086	0.245	0.234	0.268	0.23	0.233
	% coverage	90	82.297	92.828	82.687	84.15	90	86.031	98.379	84.598	85.59
BE2	5%	0.061	0.07	0.071	0.075	0.077	0.259	0.267	0.256	0.272	0.272
	95%	0.168	0.168	0.205	0.18	0.173	0.43	0.422	0.434	0.426	0.424
	90% Interval	0.107	0.098	0.134	0.105	0.096	0.171	0.155	0.178	0.154	0.152
	% coverage	90	82.444	112.332	88.489	80.177	90	81.484	93.522	80.748	79.854
BE3	5%	−0.202	−0.191	−0.203	−0.197	−0.197	−0.203	−0.19	−0.197	−0.194	−0.19
	95%	−0.079	−0.079	−0.083	−0.086	−0.086	−0.005	−0.013	−0.005	−0.014	−0.016
	90% Interval	0.123	0.112	0.12	0.111	0.111	0.198	0.177	0.192	0.18	0.175
	% coverage	90	81.686	87.374	81.248	81.175	90	80.292	87.142	81.653	79.158

Archived-t, parameter estimation from t distribution using the REML method in the archived datasets; BC, bias-corrected CI; BCa, bias correctedand accelerated CI; 90% Interval, the interval between 5% (or percentile) and 95% (or percentile); % (percent) coverage, the length of nonparametric 90% CIs measured in relation to the parametric 90% CI. (i.e., when the 2 lengths are the same, the % coverage is 90) percent of the CI which bootstrap 90% CI covers in contrast to the parametric 90% CI; %, percent for the archived datasets and percentile for the bootstrap-resampled datasets.

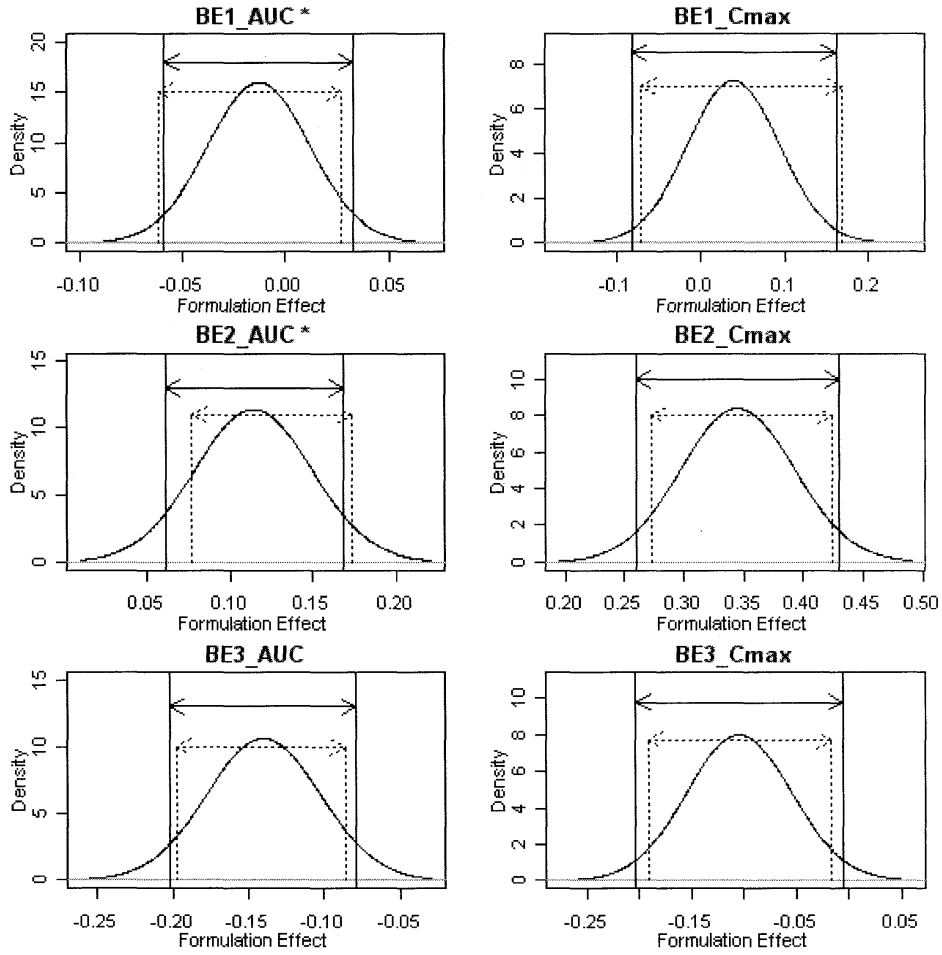

Fig. 3. Comparison of BCa 90% CIs with parametric CIs (solid lines, parametric 5% and 95% points; dotted lines, BCa 5th and 95th percentile points; density curve, anticipated parametric distribution of formulation effects from the archived datasets). *BE1_AUC and BE2_AUC were found to have non-normally distributed formulation effects by Shapro-Wilk tests.

whether samples originate from a normal distribution or not. A simple graphical way of testing normality is the normal probability (or Q-Q) plot method. A more formal test of normality is the Shapiro-Wilk test that is recommended when the sample size is small to medium. Skewness and kurtosis may also be utilized along with these tests. Unless there is direct evidence that the log-normality is not valid, log-transformed AUC and Cmax should be used for statistical analysis with the parametric method (Lacey et al., 1997). If data do not show a normal distribution even after transformation, nonparametric methods are needed. Parameters, such as Tmax that cannot be normally or log-normally distributed, also need nonparametric tests. However, when applied to a small-sized sample, as is common in BE studies, nonparametric procedures give unsatisfactory results for sensitivity. Determining the pattern of parameter distribution is a prerequisite for selecting the most appropriate statistical method for BE tests (Pabst and Jaeger, 1990; Steyn et al., 1991). In this context, we employed the bootstrap-resampling method for investigating the distribution of formulation effects. Although bootstrapped results are usually approximate, they can sometimes be more reliable and more informative than *a priori* assumptions of the distribution. The bootstrap can be implemented to inferences of correlations or ratios of variables when other analytic tools are not readily available. BE tests may be one such implementation (Sprent and Smeeton, 2001). There are several methods for calculating bootstrap CI. The percentile method is most commonly used. However, this method is highly influenced by the symmetry of the distribution pattern, and the CIs tend to be underestimated (percent coverage $< 90\%$) when the distribution of the resampled parameters is asymmetrical. The bootstrap-t method is one of the pivot methods that transform the bootstrap estimator into a pivot statistics. The distribution of pivot statistics is obtained directly from given data without the use of normal or Student's t distribution. In practice, however, bootstrap-t can be influenced by a few outliers and tends to overestimate the CIs when the distribution is skewed. The BC method corrects the asymmetry of the bootstrap distribution, but it is not a truly nonparametric method because it relies upon monotonic transformation that results in a normal distribution (Bonate, 2005). The BCa method, however, has been recommended for general use from its merit of correcting skewness (Efron and Tibshirani, 1993). In our study, the formulation effects of log (AUC) and log (Cmax) showed a normal distribution in 4 of the 6 cases. The nonparametric 90% CIs of normally distributed formulation effects were slightly different from the parametric 90% CIs of the archived datasets. However, when formulation effects were not normally distributed, the nonparametric 90% CIs based upon the bootstrap-resampling method shifted outside the parametric CIs' boundaries.

According to the EMEA guideline (EMEA, 2006), nonparametric methods are recommended for untransformed Tmax only. This is supposed to be from the fact that the number of subjects in BE studies are too small to conclude the normality of the distribution of the log-transformed AUC or Cmax by statistical tests. Nevertheless, this report exemplified the usefulness of nonparametric BE tests as an addition to the conventional BE test by comparing several different methods known.

ACKNOWLEDGEMENTS

This study was supported by a grant from the Korea Health 21 R&D Project (A060093), Ministry of Health and Welfare, Republic of Korea. A part of this manuscript was presented in poster form at the ASCPT 2007 meeting held in Orlando, U.S. With regard to the conflict of interest, the authors have nothing to declare.

REFERENCES

Bonate PL. *Pharmacokinetic-pharmacodynamic modeling and simulation.* 1st ed. Springer, New York, p 355−363, 2005.

Chow SC. *Encyclopedia of biopharmaceutical statistics.* 2nd ed. Marcel Dekker, New York, p 83−88, 2003.

Efron B, Tibshirani R. *An introduction to the bootstrap. Monographs on statistics and applied probability.* 1st ed. Chapman & Hall, New York, p 179−201, 372−391, 1993.

EMEA. *Questions & Answers on the Bioavailability and Bioequivalence Guideline (EMEA/CHMP/EWP/40326/2006)* [EMEA Web site], http://www.emea.europa.eu/pdfs/human/ewp/4032606en.pdf. European Medicines Agency, Evaluation of Medicines for Human Use. Accessed August 10, 2009.

FDA. *Guidance for industry: Statistical Approaches to Establishing Bioequivalence January 2001* [CDER Web site], http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070244.pdf. US Food and Drug Administration, Center for Drug Evaluation and Research. Accessed August 10, 2009.

Henderson AR. The bootstrap: a technique for data-driven statistics. Using computer-intensive analyses to explore experimental data. *Clin Chim Acta* 359: 1−26, 2005.

Lacey LF, Keene ON, Pritchard JF, Bye A. Common noncompartmental pharmacokinetic variables: are they normally or log-normally distributed? *J Biopharm Stat* 7: 171−178, 1997.

Midha KK, Ormsby ED, Hubbard JW, McKay G, Hawes EM, Gavalas L, McGilveray IJ. Logarithmic transformation in bioequivalence: application with two formulations of perphenazine. *J Pharm Sci* 82: 138−144, 1993.

Pabst G, Jaeger H. Review of methods and criteria for the evaluation of bioequivalence studies. *Eur J Clin Pharmacol* 38: 5−10, 1990.

Patterson SD, Jones B. Bioequivalence and the pharmaceutical industry. *Pharmaceut Statist* 1: 83−95, 2002.

Sprent P, Smeeton NC. *Applied nonparametric statistical methods.* 3rd ed. Chapman & Hall/CRC, Boca Raton, p 404−437, 2001.

Steyn HS, Koeleman HA, Gouws E, Ritschel WA. An approach to select the appropriate statistical method for testing bioequivalence. *Int J Clin Pharmacol Ther Toxicol* 29: 156−160, 1991.