CLINICAL TRIALS ARTICLE

Clinical Trials 2005; **2**: 174–181

Overlap coefficient for assessing the similarity of pharmacokinetic data between ethnically different populations

Sachiko Mizuno, Takuhiro Yamaguchi, Akira Fukushima, Yutaka Matsuyama and Yasuo Ohashi

We developed a method to assess the similarity of pharmacokinetic data between ethnically different populations. An evaluation of confidence intervals for the mean difference in pharmacokinetic parameters, such as area under the concentration-versus-time curve (AUC), between populations is often used. We propose the use of the overlap coefficient (OC), which represents the proportion of overlap between two probability distributions, as a measure of the similarity between distributions. We considered five OC estimators – two parametric ones and three nonparametric ones. Simulation studies were conducted to compare the performance of the five OC estimators and their bootstrap confidence intervals. Results showed that nonparametric estimators with fixed-bandwidth kernel density estimation had a smaller mean squared error in almost all situations, and their coverage probabilities were close to the nominal level. The proposed method was applied to pharmacokinetic data from a bridging study of a combination therapy for metastatic colorectal cancer patients in the USA and Japan. From the analyses of this study, it was suggested that the distributions of the logarithmically transformed AUC for leucovorin and 5-fluorouracil were similar between the two populations. *Clinical Trials* 2005; **2**: 174–181. www.SCTjournal.com

1 Introduction

In the development of new medical products, clinical trials are required to be standardized and conducted efficiently and quickly on a global scale. In an effort to address these issues, the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) was initiated in 1990. One outcome of the project was the development of the guideline on "Ethnic Factors in the Acceptability of Foreign Clinical Data" [1], known as the ICH E5 guideline. One of the concerns of E5 is a bridging study to allow extrapolation of foreign clinical data to a new region. The E5 guidance is based on the premise that it is not necessary to repeat the entire clinical drug development programme in the new region; bridging studies should allow for new medicines to be supplied expeditiously to patients for their benefit.

When evaluating the extrapolation of clinical data from one region to an ethnically different region, it is important to assess whether the pharmacokinetic data are similar across the populations. Pharmacokinetic studies are conducted to characterize the absorption, distribution, metabolism, and excretion of a drug either in blood or in other pertinent locations. The pharmacokinetic profiles of two populations are compared in terms of the appropriate parameters, which are measures of systemic exposure such as peak concentration and area under the concentration-versus-time curve (AUC). However, there are no standard methods for assessing the similarity of pharmacokinetic parameters. Graphical presentation of data and the evaluation of confidence intervals for the difference

Department of Biostatistics, School of Health Sciences and Nursing, University of Tokyo, Tokyo, Japan
Author for correspondence: Sachiko Mizuno, Department of Biostatistics, School of Health Sciences and Nursing, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: mizuno@epistat.m.u-tokyo.ac.jp

10.1191/1740774505cn077oa

of mean parameters between populations are often used.

The confidence interval approach claims equivalence of two populations when a two-sided confidence interval for the ratio of the geometric mean of pharmacokinetic parameters between populations is entirely contained in prespecified equivalence limits. Peace [2] argues that since the decision interval for concluding bioequivalence is based upon the ratio of means, it is more appropriate to base the decision on the mean of the individual ratios, and many statistical methods have been discussed. The method based on the ratio of the geometric mean is a basic method for testing bioequivalence [3,4]. However, there are some problems with the application of this method to a bridging study. Bioequivalence is usually assessed by comparing pharmacokinetic parameters after the administration of two formulations with the same drug product in the same individual. The assessment of bioequivalence is based on the fundamental assumption that when two formulations of the same drug product are equivalent in the rate and extent of drug absorption, they will achieve the same therapeutic effect. On the other hand, in a bridging study, the pharmacokinetic parameters of the same drug are compared between ethnically different regions. A bridging study will inevitably include patients with different baseline demographics, which may affect the profiles of the pharmacokinetic parameters. Therefore, the variability of the pharmacokinetic data in the bridging situation is essentially larger than that observed in a bioequivalence study, and it is required to show similarity rather than equivalence of the pharmacokinetic parameters.

Another problem with the confidence interval approach is that it requires equivalence with regard to the population means. The assumptions of normality and equal variances, which are often assumed in the assessment of bioequivalence, are not expected to be reasonable in the comparison of ethnically different regions. Furthermore, the methods based on a comparison of population means are, in general, statistically significant when the sample size of each group is large, even if the difference is not clinically important. In a bioequivalence study, the sample size calculation can be properly conducted using a crossover design, while it is difficult to set the sample size in the bridging settings. For these reasons, in the evaluation of the similarity of pharmacokinetic parameters across ethnically different populations, a method assessing the variability of distributions themselves is more appropriate than that based on a comparison of population means.

One approach to this problem is to measure the overlap of the distributions. The overlap coefficient (OC), the proportion of overlap of two probability distributions, has been recognized as a measure of similarity between distributions [5]. Figure 1 illustrates the OC from two normal distributions with unequal variance. The shaded area is the proportion of similar responses (OC). The OC ranges between zero and unity. The two distributions are more similar as the OC is closer to unity. Assuming that both distributions are normal and have equal variances, a comparison of two distributions using the OC approximates a comparison of two means. Rom and Hwang [6] proposed using the proportion of similar response, which is the same measure as the OC, and they applied this measure to pharmacokinetic data from a bioequivalence study under the assumption that the two distributions were normal.

The OC estimators are classified into two types according to the density estimation method: the parametric approach, with the assumption that the distributions of response are normal [6,7] and the nonparametric approach, with no assumption about the distributions [8,9]. Stine and Heyse [8] explored the properties of several estimators through simulation studies and found that the parametric approach performed well in normally distributed settings, while the nonparametric approach performed well in situations in which two samples were obtained from skewed distributions. Although the results of their study are very useful, they focused on general cases. Further examinations that reflect the bridging situation will be necessary to use OC estimates as a measure of the similarity of pharmacokinetic data between ethnically different populations.

In this paper, we propose the use of the OC for assessing the similarity of pharmacokinetic data between ethnically different populations. We consider the five OC estimators: two parametric ones and three nonparametric ones. Simulation studies

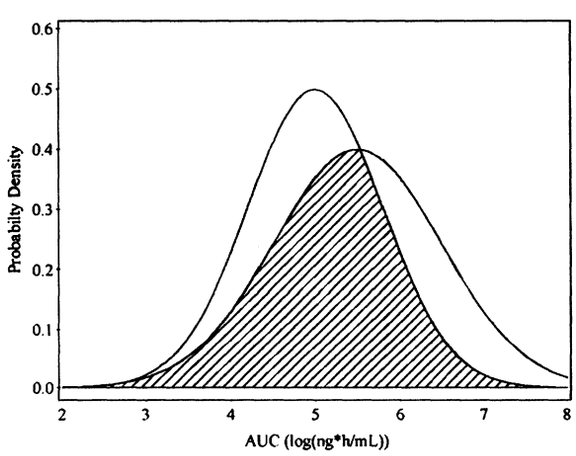

Figure 1 The overlap coefficient (OC; shaded area) from two normal distributions with unequal variance.

under various sampling situations are conducted to compare the performances of the five OC estimators and their bootstrap confidence intervals. We analyse pharmacokinetic data from a bridging study of a combination therapy for metastatic colorectal cancer patients in the USA and Japan.

2 Pharmacokinetic data of the UFT/leucovorin study

UFT is a fluorouracil prodrug and an oral 4 : 1 molar concentration of uracil plus tegafur. Combination therapy with oral UFT and oral leucovorin (UFT/leucovorin) is used to treat metastatic colorectal cancer. Multicentre phase III studies evaluating the efficacy of UFT/leucovorin have already been conducted in Europe and the USA [10,11]. To facilitate the extrapolation of these foreign clinical data to Japanese patients, pharmacokinetic studies of UFT/leucovorin for metastatic colorectal cancer patients were conducted in the USA and Japan at the same time with the same protocol [12,13]. The study subject consisted of 45 patients in the USA and 44 patients in Japan. With respect to baseline characteristics, there was a difference in the distribution of body surface area between the two populations. However, when the dosage of UFT was adjusted for body surface area, the influence of this difference was considered to be small. Blood concentrations of 5FU, uracil, and leucovorin were measured nine times within eight hours, and the AUC was calculated for each patient. We used logarithmic transformed values of the AUC for 5FU, uracil and leucovorin to assess the similarity between the two populations.

Table 1 shows the summary statistics. For all response variables, the mean values of AUC were larger in Japan than in the USA, while the standard deviations were larger in the USA than in Japan. Considering the summary statistics in Table 1, we can see that the AUC of 5FU in both populations was almost normally distributed, and little difference was observed in their variances. For uracil, the distribution of the AUC in Japan was almost normal, but that in the United States was left-skewed. For leucovorin, the distribution of the AUC was almost normal in both populations, with their variance being almost equal.

3 Estimation of the overlap coefficient

3.1 Definition of OC

Consider a pharmacokinetic study in two different populations. Let X_i $(i = 1, \ldots, n)$ and Y_i $(i = 1, \ldots, m)$ be a response variable for each patient i from two different populations with densities f and g, respectively. The OC is defined as the area under the smaller of the two population density functions:

$$OC = \int_{-\infty}^{\infty} \min[f(x), g(x)]\, dx$$

3.2 Parametric estimators

Estimation of the OC has focused on samples from a normal population. If the two populations are normal with means μ_X and μ_Y, respectively, and a common variance σ^2, the estimator of the OC is given by [5,7]

$$\widehat{OC} = 2\Phi\left(-\frac{|\hat{\delta}|}{2}\right), \qquad \hat{\delta} = \frac{\bar{X} - \bar{Y}}{S},$$

$$S = \sqrt{\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}$$

Table 1 Summary statistics of logarithmic transformed AUC for 5FU, uracil, and leucovorin

Population	Summary statistics	5FU	Uracil	Leucovorin
Japan	Sample size	44	44	44
	Mean[a]	5.21	8.54	7.80
	Standard deviation	0.64	0.59	0.41
	Skewness	0.12	−0.53	0.01
	Kurtosis	−0.12	0.49	0.19
	P-value for the test of normality[b]	0.72	0.40	0.92
United States	Sample size	43	39	42
	Mean[a]	4.82	7.78	7.62
	Standard deviation	0.83	1.03	0.43
	Skewness	−0.57	−0.76	−0.19
	Kurtosis	−0.25	0.05	−0.65
	P-value for the test of normality[b]	0.08	0.01	0.59
	P-value for the test of equal variance	0.09	0.00	0.79
	Ratio of variance	1.69	3.04	1.04

[a]log(ng.h/mL).
[b]Shapiro–Wilk test.

where $\Phi(\cdot)$ denotes the cumulative standard normal distribution, $\bar{X}$ and $\bar{Y}$ are the sample means for μ_X and μ_Y, respectively, and S_X^2 and S_Y^2 are the sample variances in each population. We refer to this estimator as "Normal-Eq".

If the two populations are normal with unequal variances and $S_X^2 < S_Y^2$, two density estimators, $\hat{f}(x)$ and $\hat{g}(x)$, cross at two points, which are given by the roots of a quadratic expression, so that:

$$(x_1, x_2) = (S_Y^2 - S_X^2)^{-1}[(\bar{X}S_Y^2 - \bar{Y}S_X^2) \pm S_X S_Y\{(\bar{X} - \bar{Y})^2 + 2(S_Y^2 - S_X^2)\log(S_X/S_Y)\}^{1/2}]$$

We take x_1 to be the smaller root. Then, the estimator of the OC is given by [5,7]

$$\widehat{OC} = 1 + \Phi(z_{11}) - \Phi(z_{12}) - \Phi(z_{21}) + \Phi(z_{22})$$

where $z_{j1} = (x_j - \bar{X})/S_X$ and $z_{j2} = (x_j - \bar{Y})/S_Y$ for $j = 1, 2$. We refer to this estimator as "Normal-Un".

3.3 Nonparametric estimators

The nonparametric estimators replace the unknown population density functions, f and g, in the OC by the kernel density estimators computed from each sample [8,9]. A kernel density is a continuous smooth estimate of the population density function. Given data X_i ($i = 1, K, n$), the kernel function $K(\cdot)$ and the bandwidth $h(X_i)$, the kernel density estimator is given by

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h(X_i)}K\left(\frac{x - X_i}{h(X_i)}\right)$$

Several simulation studies have shown that the choice of the kernel function has little effect on the estimated kernel density [8,14]. Here, we use the Gaussian kernel function, which is given by

$$K(u) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{u^2}{2}\right)$$

The bandwidth, $h(X_i)$, which is also called the smoothing parameter or window width, decides the degree of smoothness. The setting of the bandwidth is a difficult problem because a kernel density with a wide bandwidth has relatively low variance but more bias due to its wide kernel, while a small bandwidth leads to less bias but more variance. Therefore, the choice of the bandwidth strongly influences the OC estimate [8]. There are many methods for computing the bandwidth from samples, which are classified into two types: the fixed-bandwidth that has a constant value and the variable-bandwidth that changes a value of bandwidth according to data X_i adjusting for the local variations in smoothness [14].

The fixed-bandwidth is usually estimated to minimize a kernel-based estimate of the asymptotic mean integrated squared error. We introduce the notation $R(g) = \int g(u)^2\,du$ for any square integral function of $K(u)$ and $\sigma_k^2 = \int u^2 K(u)\,du$. Let f be the true density function. Under some conditions, the optimal bandwidth $h(X_i) = h_0$ is given by [14]

$$h_0 = \left[\frac{R(K)}{\sigma_k^4 R(f'')n}\right]^{1/5}$$

Here, there exist several proposals for selecting h_0 in estimating $R(f'')$. We used two methods to choose the fixed-bandwidth. The first method is the normal scale rule [15], which Stine and Heyse [8] applied to obtain nonparametric estimates of the OC. This rule depends on the assumption that the true density f is Gaussian and may tend to oversmooth if the population is multimodal. The second method is the direct plug-in rule [16], where h_0 is estimated by its sample estimate of $R(f'')$. We refer to the OC estimators based on these two fixed-bandwidth kernel estimators as "Kernel-Ns" and "Kernel-Pi", respectively.

The usual fixed-bandwidth density estimator is susceptible to bumpiness in the tails, since it does not adapt to local variations in the smoothness. The estimator can be generalized to allow this by using the adaptive bandwidth. In this variable kernel density estimator, the $h(X_i)$ is replaced by its sample values. In this study, we used the variable-bandwidth $h(X_i) = h_0 f(X_i)^{-1/2}$, which was recommended by Abramson [17]. We refer to the OC estimator based on the variable-bandwidth kernel estimator as "Kernel-Var".

3.4 Bootstrap confidence intervals

The confidence intervals can be easily constructed for Normal-Eq and Normal-Un estimators [18]. However, it is difficult to estimate the standard error of the nonparametric estimators. We consider bootstrap confidence intervals for the five estimators (Normal-Eq, Normal-Un, Kernel-Ns, Kernel-Pi, and Kernel-Var). We used a standard normal interval, which is the simple bootstrap interval based on a normal approximation [19]. The implementation is an iteration of the steps used to estimate the OC. Each iteration of the resampling procedure is as follows:

1) Draw samples with replacement from two observed groups with size n and m, respectively. The bootstrap samples are of the same size as the initial samples.
2) For the nonparametric estimators, compute the kernel estimates of population density.
3) Compute the estimates of the OC from the bootstrap samples or the bootstrap kernel estimates calculated in step 2.

This procedure is repeated B times. Then, the bootstrap estimate of standard error SE^* is given by

$$SE^* = \sqrt{\frac{\sum_{b=1}^{B} (OC_b^* - \overline{OC^*})^2}{B-1}}$$

where OC_b^* is the estimate from the bth sample ($b = 1,\ldots, B$) and $\overline{OC^*} = \sum_{b=1}^{B} OC_b^*/B$. The standard normal interval is $\widehat{OC} \pm z_{1-\alpha/2}SE^*$, where $z_{1-\alpha/2}$ is the $1-\alpha/2$ percentile of the standard normal distribution. For our simulations and analyses of actual data, we set $B = 200$.

All analyses were performed using the SAS software package, IML procedure, QUAD subroutine [20]. This subroutine can implement an adaptive global-type integrator that produces a quick, rough estimate of the integration result and then refines the estimate until achieving the prescribed accuracy.

4 Simulation study

To evaluate the performance of five estimators of the OC, we carried out simulation studies to approximate the AUC data in the UFT/leucovorin study. The following three situations were considered for the shapes of the distributions of two populations. Situation 1 was based on the leucovorin data; both distributions were normal, with equal variance. Situation 2 was based on the 5FU data; both distributions were normal, with unequal variance. Situation 3 was based on the uracil data; one distribution was normal, and the other was not normal. The skewed distribution was simulated by a mixture of two normal distributions considering the standard deviation, skewness and kurtosis in the uracil data. In the bridging situation, it is unlikely that the true value of the OC is almost 1, and the two distributions are not considered to be similar if the OC is less than 0.6. Therefore, the mean differences of the two populations Δ were set so that the true values of the OC ranged from 0.6 to 0.9. The detailed simulation settings in each situation are as follows:

Situation 1

Let $N(\mu, \sigma^2)$ be the normal distribution with mean μ and variance σ^2, and let X_i and Y_i be the independent samples from each population.

$$X_i \sim N(\Delta, 1^2), \qquad Y_i \sim N(0, 1^2)$$

where $\Delta = 0.25$, 0.50, 0.75, and 1.00. In this case, the true values of the OC were 0.90, 0.80, 0.71, and 0.62, respectively.

Situation 2

$$X_i \sim N(\Delta, 1^2), \qquad Y_i \sim N(0, (1/\rho)^2)$$

where the variance ratio ρ^2 was 2 and $\Delta = 0$, 0.50, and 0.80. In this case, the true values of the OC were 0.82, 0.71, and 0.60, respectively.

Situation 3

$$X_i \sim N(\Delta, 1^2), \qquad Y_i \sim pN(0, 1^2) + (1-p)N(-1.8, 1.4^2)$$

where the mixture proportion $p = 0.6$ and $\Delta = 0$, 0.30, and 0.60. In this case, the true values of the OC were 0.78, 0.70, and 0.61, respectively.

An equal sample size of 50 for each group was generated, and the simulations were based on 1000 replications, so that the estimated coverage probabilities of true 90% confidence intervals would have a simulation accuracy of approximately $\pm 1.6\%$ due to Monte Carlo variability. The simulations were evaluated in terms of the mean squared error (MSE), the coverage probability of 90% confidence intervals, and the mean length of 90% confidence intervals.

Table 2 shows the simulation results for the MSE. In situation 1, Kernel-Var had the largest MSE when $\Delta = 1$. There were no large differences among the other estimators. In situation 2, Normal-Eq had the largest MSE for all values of Δ. The MSE of Normal-Un, Kernel-Ns, Kernel-Pi, and Kernel-Var were similar, with Kernel-Var having the largest. In situation 3, Normal-Un had the largest MSE for all values of Δ, in particular, a large MSE when $\Delta = 0$ and 0.3. Among nonparametric estimators, Kernel-Var had the largest MSE.

Table 3 shows the simulation results for the coverage probability and mean length of 90% confidence intervals. The coverage probabilities were very close to the nominal level of 90% for three nonparametric estimators (Kernel-Ns, Kernel-Pi, and Kernel-Var) in all situations. On the other hand, the coverage probabilities for two parametric estimators (Normal-Eq and Normal-Un) were smaller than the nominal level of 90% in situation 2 and situation 3, respectively. In terms of the confidence interval length, no large differences were observed among estimators, although the lengths of Normal-Eq were a little wider than those of the other estimators.

The results of the simulation studies are summarized as follows. Normal-Eq had the largest MSE with poor coverage probabilities in situation 2. Normal-Un had the largest MSE with poor coverage probabilities in situation 3. In almost all cases, Kernel-Var had a larger MSE than the other nonparametric estimators (Kernel-Ns and Kernel-Pi).

Table 2 Mean squared error (MSE) of the OC estimates

	True value		Estimators[b]				
Situation[a]	Δ	OC	Normal-Eq	Normal-Un	Kernel-Ns	Kernel-Pi	Kernel-Var
Situation 1	0.25	0.90	0.0047	0.0047	0.0049	0.0050	0.0044
	0.50	0.80	0.0062	0.0058	0.0048	0.0047	0.0044
	0.75	0.71	0.0058	0.0056	0.0053	0.0052	0.0055
	1.00	0.62	0.0057	0.0056	0.0059	0.0059	0.0076
Situation 2	0.00	0.82	0.0139	0.0042	0.0049	0.0044	0.0060
	0.50	0.71	0.0082	0.0048	0.0048	0.0047	0.0052
	0.80	0.60	0.0062	0.0048	0.0048	0.0049	0.0054
Situation 3	0.00	0.78	0.0052	0.0085	0.0050	0.0047	0.0056
	0.30	0.70	0.0055	0.0084	0.0046	0.0046	0.0048
	0.60	0.61	0.0049	0.0066	0.0050	0.0051	0.0062

[a]Situation 1: two equal-variance normal distributions; situation 2: two unequal-variance normal distributions; situation 3: normal distribution and skewed distribution (mixture distribution).
[b]Normal-Eq: parametric approach (equal-variance); Normal-Un: parametric approach (unequal-variance); Kernel-Ns: nonparametric approach (fixed-bandwidth, normal scale rule); Kernel-Pi: nonparametric approach (fixed-bandwidth, plug-in rule); Kernel-Var: nonparametric approach (variable-bandwidth).

5 Results of the UFT/leucovorin study

From the results of the simulation studies, the Kernel-Ns and the Kernel-Pi outperformed the other estimators in terms of the MSE, the coverage probability and the mean length of the 90% confidence intervals. Therefore, we used these two estimators for the analysis of the UFT/leucovorin pharmacokinetic data. We estimated the OC and its 90% confidence intervals of logarithmic transformed AUC for 5FU, uracil, and leucovorin. We also estimated the ratio of the geometric mean and its 90% confidence intervals, which is often used for the assessment of bioequivalence.

Table 4 shows the results from these analyses. Few differences in the two OC estimators were observed, as expected. The OC estimate for leucovorin had the largest value, and the OC estimate for 5FU was close to that for leucovorin. On the other hand, the OC estimate for uracil was a relatively small value. For the ratio of the geometric mean, the estimate for leucovorin had a value close to 1, although the estimates for 5FU and uracil were substantially larger than 1.

6 Discussion

We compared the relative performance of the five OC estimators and their bootstrap confidence intervals in bridging studies where the similarity of pharmacokinetic profiles was assessed between ethnically different populations. The simulations have been conducted so as to approximate the AUC data in the UFT/leucovorin study. From the simulation studies, although the simulations were limited, we can draw some conclusions about the performance of the five estimators. First, nonparametric methods outperformed parametric methods in almost all situations. Among nonparametric methods, Kernel-Ns and Kernel-Pi gave consistently good results, and few differences in their performances were observed. Therefore, it is suggested that these two nonparametric estimators with fixed-bandwidth kernel are robust to the distributional assumptions and provide a reasonably good measure of similarity in the analysis of pharmacokinetic data in bridging studies. For selecting bandwidth, we used two simple methods, the normal scale rule (Kernel-Ns) [15] and the direct plug-in rule (Kernel-Pi) [16]. Sheather and Jones [21] have proposed a more theoretically attractive method for selecting the bandwidth in kernel density estimation. However, their approach needs to use an optimization technique such as the Newton–Raphson method to obtain the bandwidth. Considering the stability and simplicity for estimating the bandwidth, Kernel-Ns and Kernel-Pi seem to be sufficient to estimate the OC in practical settings.

Secondly, for the parametric estimators, Normal-Eq and Normal-Un provided poor performance under the normal distribution with unequal variance (situation 2) and under the non-normal distributions (situation 3), respectively. These findings are compatible with the simulation results of Stine and Heyse [8]. However, the MSE of Normal-Eq in situation 3 was not as large as one might have expected. This relatively good performance of the Normal-Eq appears to be related to the simulation settings of the skewed distribution. In our simulations, the skewed distribution was simulated by a mixture of two normal distributions, which was not so heavily skewed. In the linear discriminant analysis, it is well known that the estimation of

Table 3 Coverage probability and mean length of nominal 90% confidence intervals of the OC estimates

	True value		Estimators[a]									
			Normal-Eq		Normal-Un		Kernel-Ns		Kernel-Pi		Kernel-Var	
Situation[b]	Δ	OC	Coverage probability	Length[c]	Coverage probability	Length[c]	Coverage probability	Length[c]	Coverage probability	Length[c]	Coverage probability	Length[c]
Situation 1	0.25	0.90	0.85	0.22	0.90	0.21	0.87	0.19	0.86	0.20	0.90	0.21
	0.50	0.80	0.87	0.25	0.88	0.23	0.91	0.22	0.90	0.21	0.88	0.22
	0.75	0.71	0.88	0.25	0.87	0.24	0.87	0.23	0.87	0.23	0.87	0.23
	1.00	0.62	0.88	0.25	0.87	0.24	0.85	0.24	0.86	0.23	0.82	0.24
Situation 2	0.00	0.82	0.12	0.20	0.87	0.20	0.90	0.22	0.90	0.21	0.87	0.23
	0.50	0.71	0.60	0.25	0.86	0.21	0.85	0.22	0.85	0.22	0.85	0.23
	0.80	0.60	0.82	0.25	0.86	0.21	0.87	0.22	0.88	0.22	0.89	0.23
Situation 3	0.00	0.78	0.88	0.24	0.75	0.23	0.88	0.21	0.90	0.21	0.87	0.23
	0.30	0.70	0.90	0.23	0.76	0.23	0.87	0.22	0.88	0.22	0.86	0.23
	0.60	0.61	0.89	0.22	0.80	0.22	0.88	0.22	0.88	0.23	0.86	0.23

[a]Situation 1: two equal-variance normal distributions; situation 2: two unequal-variance normal distributions; situation 3: normal distribution and skewed distribution (mixture distribution).
[b]Normal-Eq: parametric approach (equal-variance); Normal-Un: parametric approach (unequal-variance); Kernel-Ns: nonparametric approach (fixed-bandwidth, normal scale rule); Kernel-Pi: nonparametric approach (fixed-bandwidth, plug-in rule); Kernel-Var: nonparametric approach (variable-bandwidth).
[c]Mean confidence interval length.

false discriminant rates is robust to non-normality, such as mixtures of normal distributions [22,23].

From the analysis of the UFT/leucovorin study, we tried to judge the similarity of the AUCs of 5FU, uracil, and leucovorin between the two populations. In a bioequivalence study, assessment for equivalence of AUC between the two populations is based on whether 90% confidence intervals for the ratio of geometric means are entirely contained in the range from 0.8 to 1.25 [3,4]. According to this criterion, all response variables are judged to be dissimilar between the two populations. However, this criterion is considered to be strict in the bridging situation, because of the differences in study design. For the evaluation of bioequivalence of the maximum plasma concentration (C_{max}), a wider interval from 0.7 to 1.43 may be acceptable [24]. It is reasonable to suppose that this wider criterion will suffice for the assessment of similarity in the bridging situation. According to this criterion, only leucovorin data are judged to be similar between the two populations.

There is no general consensus about the criteria of similarity by the OC estimates. However, considering the relationship between the standardized mean difference and the OC estimates in the normal distributions with equal variance, judgement using the confidence intervals for the difference of the population means is equivalent to the judgement by the lower limit of the OC estimates. In leucovorin, the distributions of AUC were almost normal in both populations, with their variances being almost equal. The lower limits of the 90% confidence interval for two nonparametric OC estimates (Kernel-Ns and Kernel-Pi) were 0.75 and 0.76, respectively. Therefore, it seems reasonable to say that the two distributions of AUC for leucovorin are similar.

For 5FU, the distributions of AUC were almost normal, and large differences were not observed in their variances. Furthermore, there was little difference in the lower limits of the OC estimates between leucovorin and 5FU. It can be interpreted that the two distributions of AUC for 5FU are similar to some extent.

For uracil, although the distribution of AUC in the USA was left-skewed, the lower limits of the OC estimates were smaller than those of leucovorin. It may be interpreted that the two distributions of AUC for uracil are not similar. With respect to criteria to assess similarity using the OC estimates, further research and considerable effort to reach the consensus among specialists, including the regulatory authority, will be needed.

Considering the low values of the OC estimates for uracil, the presence of ethnic factors that influence the pharmacokinetic profile of uracil may be suggested. It is important to examine the

Table 4 Estimates and 90% confidence intervals of the OC and the ratio of geometric mean

Methods[a]		Estimates (90% confidence intervals)		
		5FU	Uracil	Leucovorin
OC	Kernel-Ns	0.82 (0.72, 0.92)	0.67 (0.55, 0.78)	0.85 (0.75, 0.94)
	Kernel-Pi	0.83 (0.73, 0.93)	0.69 (0.58, 0.79)	0.86 (0.76, 0.95)
Ratio of geometric mean		1.47 (0.92, 2.02)	2.13 (1.48, 3.06)	1.19 (0.99, 1.42)

[a] Kernel-Ns and Kernel-Pi: nonparametric approaches with kernel density estimation using different fixed-bandwidths (normal scale rule and plug-in rule) for the overlap coefficient (OC).

influence of ethnic factors, such as weight and body surface area, on pharmacokinetic profiles between patients in different regions. If some ethnic factors are known to influence the pharmacokinetic data, the statistical methods that adjust such factors as covariates will be applied. One of those methods is the analysis of covariance, which is based on mean differences between two data sets. With respect to the OC estimates, there is no general methodology for such analyses. Further research on this issue will be necessary.

Acknowledgements

We are grateful to Dr. Hyodo and many other participants in the UFT/leucovorin study for providing valuable data.

References

1. **ICH steering committee.** International Conference on Harmonisation (ICH) harmonized tripartite guideline. Ethnic factors in the acceptability of foreign clinical data (E5). *Federal Register, Food and Drug Administration* 10 June 1998; **63**(111): 31790.
2. **Peace KE.** Estimating the degree of equivalence or non-equivalence: An alternative to bioequivalency testing. *American Statistical Association, Proceedings of the Biopharmaceutical Section* 1986; 63–69.
3. **US Food and Drug Administration.** Guidance on statistical procedures for bioequivalence using a standard two-treatment crossover design. Rockville, MD: Division of Bioequivalence, Office of Generic Drugs, Center for Drug Evaluation and Research, Food and Drug Administration, 1992.
4. **US Food and Drug Administration.** Guidance for industry: Statistical approaches to establishing bioequivalence. Rockville, MD: Division of Communications Management, Office of Training and Communications, Center for Drug Evaluation and Research, Food and Drug Administration, 2001.
5. **Bradley EL.** Overlapping coefficient. In Kotz S, Johnson NL, Read CB, eds. *Encyclopedia of statistical science.* New York: Wiley, 1985; **6**: 546–47.
6. **Rom DM, Hwang E.** Testing for individual and population equivalence based on the proportion of similar responses. *Stat Med* 1996; **15**: 1489–505.
7. **Inman HF, Bradley EL.** The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. *Commun Statist Theory Methods* 1989; **18**: 3851–74.
8. **Stine RA, Heyse JF.** Non-parametric estimates of overlap. *Stat Med* 2001; **20**: 215–36.
9. **Clemons TE, Bradley EL.** A nonparametric measure of the overlapping coefficient. *Comput Statist Data Anal* 2000; **34**: 51–61.
10. **Pazdur R, Douillad JY, Skillings JR** *et al.* Multicenter phase III study of 5-fluorouracil (5-FU) or UFT in combination with leucovorin (LV) in patients with metastatic colorectal cancer. *Proc Am Soc Clin Oncol* 1999; **18**: 263a.
11. **Carmichael J, Popiela T, Randstone D** *et al.* Randomized comparative study of ORZEL (Oral Uracil/Tegafur (UFT) plus leucovorin (LV)) versus parenteral 5-fluorouracil (5-FU) plus LV in patients with metastatic colorectal cancer. *Proc Am Soc Clin Oncol* 1999; **18**: 246a.
12. **Hoff PM, Loehrer P, Abbruzzese JL** *et al.* Joint United States of America (USA)/Japan study of UFT (uracil and tegafur) plus leucovorin (LV) in patients (pts) with metastatic colorectal cancer. *Eur J Cancer* 2001; **37**(suppl. 6): s307.
13. **Hyodo I, Ohtsu A, Shirao K.** Joint United States of America (USA)/Japan study of oral Uracil/Ftorafur (UFT) plus leucovorin in patients with metastatic colorectal carcinoma. *J Japan Soc Cancer Ther* 2001; **36**: 538.
14. **Simonoff JS.** *Smoothing methods in statistics.* New York: Springer-Verlag, 1996.
15. **Silverman BW.** *Density estimation for statistics and data analysis.* London: Chapman and Hall, 1986.
16. **Wand MP, Jones MC.** *Kernel smoothing.* London: Chapman and Hall, 1995.
17. **Abramson IS.** On bandwidth variation in kernel estimates: A square root law. *Annal Stat* 1982; **10**: 1217–23.
18. **Inman HF, Bradley EL.** Hypothesis tests and confidence interval estimates for the overlap of two normal distributions with equal variances. *Environmetrics* 1994; **5**: 167–89.
19. **Efron B, Tibshirani RJ.** *An introduction to the bootstrap.* New York: Chapman and Hall, 1993.
20. **SAS.** *SAS/IML user's guide. Release 6.07.* Cary, NC: SAS Institute Inc., 1990.
21. **Sheather SJ, Jones MC.** A reliable data-based bandwidth selection method for kernel density estimation. *J Roy Stat Soc B* 1991; **53**: 683–90.
22. **Balakrishnan N, Kocherlakota S.** Robustness to nonnormality of the linear discriminant function: mixtures of normal distributions. *Commun Statist Theory Methods* 1985; **14**: 465–78.
23. **Reiser B.** Confidence intervals for the mahalanobis distance. *Commun Statist Simul Comput* 2001; **30**: 37–45.
24. **Committee for Proprietary Medicinal Products.** Note for guidance on the investigation of bioavailability and bioequivalence. London: The European Agency for the Evaluation of Medicinal Products, 26 July 2001.