Statistica Sinica **9**(1999), 1011-1027

STATISTICAL EVALUATIONS OF DISSOLUTION SIMILARITY

Mi-chia Ma*, Ru-pyng Lin* and Jen-pei Liu*†

National Cheng-kung University and †National Health Research Institutes

Abstract: Statistical properties of several criteria for assessment of similarity between two dissolution profiles are investigated. These include the similarity factor f_2, and metrics based on the mean squared distance and the mean absolute difference. The probability density function of f_2 and its first two moments are derived under the assumption of multivariate normality, with special attention to compound symmetry covariance structure. The intractable nature of the distribution of f_2 is demonstrated. Empirical results from a large simulation study are also presented. Advantages and drawbacks of proceduces based on the mean absolute difference and mean squared distance are discussed.

Key words and phrases: Dissolution, similarity factor.

1. Introduction

The U.S. Food and Drug Administration (FDA) has recently encouraged pharmaceutical companies to explore the relationship between *in vivo* drug bioavailability and *in vitro* dissolution. The *in vivo* bioavailability study is to investigate the rate and extent of drug absorption in humans. On the other hand, drug absorption depends upon the dissolved state of the drug product. Lesson (1995) suggested that *in vitro* dissolution testing be used as a surrogate for *in vivo* bioequivalence studies to assess equivalence between the test and reference formulations, and for postapproval changes.

On November 30, 1995, the FDA issued a guidance *Immediate Release Solid Oral Dosage Forms; Scale-up and Postapproval Changes: Chemistry, Manufacturing, and Controls; In Vitro Dissolution Testing; In Vivo Bioequivalence Documentation* (*SUPAC-IR*) (Federal Register, Vol. 60, No. 230, Notices, PP. 61638-61643). The *SUPAC-IR* provides recommendations to sponsors of new drug applications (NDA's), abbreviated new drug applications (ANDA's), and abbreviated antibiotic applications (AADA's) who intend, during the postapproval period, to change (i) the components or compositions; (ii) the site of manufacture; (iii) the scale-up/scale-down of manufacture; and/or (iv) the manufacturing (process and equipment) of an immediate release oral formulation. For each type of change, the *SUPAC-IR* also defines (i) levels of changes; (ii) recommended

chemistry, manufacturing, and controls tests for each level of change; (iii) *in vitro* dissolution and/or *in vivo* bioequivalence tests for each level of change; and (iv) documentation that should support the change.

If dissolution profile similarity is demonstrated for the formulations before and after the changes, then expensive *in vivo* bioequivalence testing can be waived. Dissolution testing should be conducted on at least 12 dosage units for both formulations. Other FDA guidances on dissolution testing include dissolution testing of immediate release solid oral dosage forms (1997), extended release solid oral dosage forms (1997), and *SUPAC-MR* (1997).

Various procedures have been proposed for statistical assessment of dissolution profile similarity (Chow and Ki (1997), Sathe, Tsong and Shah (1996a, b), Tsong, Hammerstrom, Sathe and Shah (1996, 1997), Polli, Rekhi, Augsburger and Shah (1997) and O'Hara, Dunne, Kinahan, Cunningham, Stark and Devane (1996)). These methods include application of either a nested model (Gill (1988)) or an autoregressive time series model (Chow and Ki (1997)) to the correlations between cumulative percents dissolved at different time points, and consideration of Mahalanobis distance (Tsong, Hammerstrom, Sathe and Shah (1996)) as a criterion for assessment of similarity in dissolution profiles between two formulations.

In the FDA guidance on *Dissolution Testing for Immediate Release Solid Dosage Form* (1997), three statistical methods are suggested for evaluation of dissolution profile similarity. These are Mahalanobis distance, model-fitting, and the similarity factor f_2 proposed by Moore and Flanner (1996). The methods of Mahalanobis and model-fitting have been reviewed and discussed (Tsong, Hammerstrom, Sathe and Shah (1996), Sathe, Tsong and Shah (1996a)). In addition, the *SUPAC-IR* uses f_2 to compare the dissolution profiles between the test and reference formulations. The *SUPAC-IR* suggests that two dissolution profiles are similar if f_2 is between 50 and 100. Although f_2 is easy to implement, there are statistical issues which need to be clarified and resolved for its regulatory implementation (Liu, Ma and Chow (1997)). Recently, other criteria such as mean absolute distance and mean squared distance have been proposed for evaluation of dissolution profile similarity (Tsong, Hammerstrom, Sathe and Shah (1996)). However, literature on their estimators and the corresponding properties is scant. In Section 2, the formal definition of the similarity factor f_2 specified in the *SUPAC-IR* is given. The probability density function and asymptotic properties of f_2 are also derived in Section 2. The simulation results of f_2 are presented in Section 3. Estimators for the mean absolute difference and mean squared distance are given in Section 4. Final remarks are provided in Section 5.

2. Distribution of the Similarity Factor f_2

Let X_{jki} be the observed cumulative percent dissolved for dosage unit i at sampling time point k for formulation j, where $j = R, T$, $i = 1, \ldots, I$, $k = 1, \ldots, n$, and let R and T denote the reference and test formulation, respectively. Let W be the sum of squares of differences in average cumulative percent dissolved between the reference and test formulations over all sampling time points, $W = \sum_{k=1}^{n} (\bar{X}_{Tk.} - \bar{X}_{Rk.})^2$, where $\bar{X}_{jk.}$ is the average cumulative percent dissolved at all sampling time point k for formulation j, $j = R, T$, $k = 1, \ldots, n$. The statistic W is a natural summary quantity measuring the overall closeness between two dissolution profiles. However, it is not properly scaled for easy and routine application by chemists or pharmacists. As a result, the similarity factor f_2 was proposed (Moore and Flanner (1996)). It can be expressed in terms of W as

$$f_2 = 50 \times \log_{10} \left[\left(1 + \frac{W}{n} \right)^{-\frac{1}{2}} \times 100 \right] = 100 - \frac{25}{\ln 10} \ln(1 + W/n). \tag{2.1}$$

Note that the cumulative percent dissolved is bounded between 0 and 100. It follows that the difference in average cumulative percent dissolved is also bounded between 0 and 100. As a result f_2 ranges approximately from 0 to 100. The *SUPAC-IR* suggested that similarity between two dissolution profiles be concluded if f_2 is between 50 and 100, where 50 represents an average 10% difference at all sampling time points and 100 is the upper bound of f_2 when the distance at all sampling time points is 0.

Let $\underline{D} = (\bar{X}_{T1.} - \bar{X}_{R1.}, \ldots, \bar{X}_{Tn.} - \bar{X}_{Rn.})'$. Assume that $\underline{D} \sim N(\underline{\mu}_D, \Sigma)$, where $\underline{\mu}_D = \underline{\mu}_T - \underline{\mu}_R = (\mu_{D1}, \ldots, \mu_{Dn})'$ and Σ is an $n \times n$ positive definite covariance matrix. Since $W = \underline{\mathrm{D}}'\underline{\mathrm{D}}$ we have

$$E(W/n) = \underline{\mu}_D' \underline{\mu}_D / n + tr[\Sigma]/n = \mu_D^2 + \sigma_D^2, \tag{2.2}$$

where μ_D^2 and σ_D^2 denote the mean square in the population average differences and the mean of variances over all sampling time points, respectively.

The probability density function of f_2 is given by

$$\begin{aligned} f(v) = & \int_0^\infty \cdots \int_0^\infty \prod_{i=1}^{n-1} \left[\sum_{k=0}^{\infty} g_k(\tau_i) f_{1+2k}(v_i) \right] \\ & \times \sum_{k=0}^{\infty} g_k(\tau_n) f_{1+2k} \left[\frac{n \times \left(10^{\frac{100-v}{25}} - 1 \right) - \sum_{i=1}^{n-1} \lambda_i v_i}{\lambda_n} \right] \times \frac{1}{\lambda_n} \\ & \times \left| \frac{n \times \ln(10)}{25} \times 10^{\frac{100-v}{25}} \right| dv_1 \cdots dv_{n-1}, \end{aligned} \tag{2.3}$$

where $g_k(\tau_i) = e^{-\tau_i/2} \times \frac{(\tau_i/2)^k}{k!}$, f_{1+2k} is the density function of the chi-square distribution with $1+2k$ d.f., $\tau_i = \left(\underline{\mu}'_D \underline{e}_i\right)^2 / \lambda_i$, and λ_i and $\underline{e}_i$ are the eigenvalues and eigenvectors of Σ, respectively.

Under the assumption of compound symmetry for Σ, the expected value and variance of W are given by

$$E(W) = \sum_{k=1}^{n} \mu_{Dk}^2 + \frac{2n}{I}(\sigma_e^2 + \sigma_s^2), \tag{2.4}$$

$$\text{Var}\,(W) = \frac{8}{I}[\sigma_s^2(\sum_{k=1}^{n} \mu_{Dk})^2 + \sigma_e^2 \sum_{k=1}^{n} \mu_{Dk}^2] + \frac{8n}{I^2}(\sigma_e^2 + \sigma_s^2)^2[1 + (n-1)K^2], \tag{2.5}$$

where $K = \frac{\sigma_s^2}{\sigma_e^2 + \sigma_s^2}$, $\sigma_e^2 + \sigma_s^2$ is the variability of the cumulative dissolved observed at a sampling time point, and σ_s^2 represents the covariance between different sampling time points.

Even assuming normality, (2.3) is very complicated. Failure to find the closed form of the expected value of f_2 prevents us from constructing confidence interval by any means including bootstrap confidence interval since we do not know the parameter that the confidence interval is constructed for. If $W/n \leq 1$, the expected value and variance of f_2 may be approximated by a Taylor series expansion about $E(W)$:

$$E(f_2) \approx 100 - \frac{25}{\ln 10} \ln[1 + E(\frac{W}{n})] + \frac{25}{2 \ln 10}[\frac{1}{n + E(W)}]^2 \text{Var}\,(W), \tag{2.6}$$

$$\begin{aligned} \text{Var}\,(f_2) \approx\ & [\frac{25}{(\ln 10)(n + E(W))}]^2 \text{Var}\,(W) \\ & + \frac{1}{4}\{\frac{25}{(\ln 10)[n + E(W)]^2}\}^2 \{E[W - E(W)]^4 - [\text{Var}\,(W)]^2\} \\ & - (\frac{25}{\ln 10})^2 \frac{1}{[n + E(W)]^3} E[W - E(W)]^3. \end{aligned} \tag{2.7}$$

To utilize f_2, we need to find the population parameters that the similarity factor tries to estimate or test. From (2.6), if we only take the first two terms of the approximate expected value of f_2, $E(f_2) \approx 100 - \frac{25}{\ln 10} \ln[1 + \mu_D^2 + \sigma_D^2]$. Then the possible range of $E(f_2)$ is from 0 to 100, and the upper limit of 100 is reached when both μ_D^2 and σ_D^2 are equal to 0.

The expected value of f_2 may also be approximated by a Taylor series expansion about $E(\underline{D}) = \underline{\mu}_D$:

$$\theta = 100 - \frac{25}{\ln 10} \ln[1 + \mu_D^2]. \tag{2.8}$$

Note that $\theta \geq 100 - \frac{25}{\ln 10} \ln[1 + \mu_D^2 + \sigma_D^2]$. Thus if one considers θ as the criterion, application of f_2 to the assessment of dissolution profile similarity will be conservative.

From (2.6), $E(f_2)$ is a function of $E(W/n)$. As a result, two dissolution profiles are concluded similar if $E(W/n)$ is smaller than some pre-specified allowable upper limit. It follows that the statistical hypotheses corresponding to the similarity between dissolution profiles may equivalently be formulated as a one-sided hypothesis in terms of mean square of population average differences and mean variability as

$$H_0 : \mu_D^2 + \sigma_D^2 \geq \Delta \qquad v.s. \qquad H_a : \mu_D^2 + \sigma_D^2 < \Delta, \tag{2.9}$$

where Δ is some positive allowable upper limit for concluding dissolution similarity.

From (2.9), inference about the similarity factor f_2 is for population average differences and population variances jointly. The joint statistical inference for average and variability can be very difficult since more unknown parameters are involved. We can not make inference to the individual components based on f_2 without additional assumptions about unknown parameters. For comments regarding f_2 for the assessment of dissolution similarity, see Liu, Ma and Chow (1997).

From (2.6) and (2.7), if we only take the first two terms of the expansion of f_2 under $\underline{\mu}_D = 0$, $\sigma_s^2 + \sigma_e^2 = 27.6312$, $I = 12, n = 7$ and $\sigma_s^2 = 2.76312$, the approximate expected value of f_2 is about 81.285, and the first term of approximate variance of f_2 is about 24.099. But, if we take the first three terms of the expansion of f_2 under the same condition, then the approximate expected value of f_2 is about 82.395 and the second term and the third term of approximate variance of f_2 are about 5.2775 and 12.9148, respectively. As a result, the approximate variance of f_2 is about 16.462. The approximate variance of f_2 therefore fluctuates more than the approximate mean of f_2.

3. Simulation on Sample Statistic of f_2

We conducted a simulation to investigate the distribution of f_2. The assumptions for our simulation are given below.

(1) The differences in the average cumulative percent dissolved between the test and reference formulation follow a multivariate normal distribution.
(2) The differences in the average cumulative percent dissolved are constant at all sampling time points.
(3) The covariance matrices of test and reference formulations have the same compound symmetry structure.

Under compound symmetry, W can be expressed as $W = \lambda_1 Y_1 + \lambda_2 Y_2$, where Y_1 and Y_2 are independent, Y_1 follows a noncentral chi-square distribution with 1 d.f. and noncentral parameter $\tau_1 = (\underline{\mu}'_D \underline{e}_n)^2/\lambda_1$, $\lambda_1 = \frac{2}{I}(\sigma_e^2 + n\sigma_s^2)$, and Y_2 follows a noncentral chi-square distribution with $(n-1)$ d.f. and noncentral parameter $\tau_2 = \underline{\mu}'_D \Sigma^{-1} \underline{\mu}_D - \tau_1$, $\lambda_2 = \frac{2\sigma_e^2}{I}$. If $\underline{\mu}_D = \underline{0}$ then $\tau_1 = 0$ and $\tau_2 = 0$. We invoked the RANGAM function of SAS version 6.06 on a DEC work station for independent central chi-square random variables when the $\underline{\mu}_D = \underline{0}$. If $\underline{\mu}_D = 3\underline{1}$ then $\tau_2 = 0$, where $\underline{1}$ is a $n \times 1$ unit vector. We used the SAS function RANUNI and CINV to generate random samples when the $\underline{\mu}_D = 3\underline{1}$. The similarity factor f_2 was then calculated for each sample. Our simulation explored 5 magnitudes of total variability: 27.6312, 138.156, 276.312, 552.624 and 594. For each total variability, the ratio of inter-unit variability to the total variability, denoted by K, was chosen from 0.1 to 0.9 by an interval of 0.1. Four different sample sizes of 6, 12, 18 and 24 units were considered for each formulation. Because the number of sampling time points for a dissolution testing is usually greater than 2, we considered 5 different numbers of sampling time points: 3, 4, 5, 6 and 7 in this simulation. Because f_2 is invariant to unequal spacing between sampling time points, we do not have to specify the length of interval between time points. The mean, variance, the lower 5 % quantile of the empirical distribution, and the empirical probability of concluding similarity according to the criterion based on f_2 specified in the *SUPAC-IR* were obtained based on ten thousand (10,000) random samples for each of 900 combinations. Because the results are consistent, we only present those for the combination $\underline{\mu}_D = \underline{0}$ and $\underline{\mu}_D = 3\underline{1}$, total variability=27.6312, $K =$ 0.1 to 0.9 by 0.2; $I = 6$ to 24 by 6; and $n = 3$ to 7 by 2. Other results can be obtained from the authors upon request. The empirical mean and variance are given in Tables 1 and 2 for $\underline{\mu}_D = \underline{0}$ and $\underline{\mu}_D = 3\underline{1}$, respectively, with the approximate mean computed by (2.6) and the approximate variance by the first term of (2.7). The results of simulation are summarized below.

(1) Figure 1 displays the histogram of the empirical distribution of f_2 for $I = 12$, $n = 7$, $K = 0.5$, $\underline{\mu}_D = \underline{0}$, and total variability 27.6312. From Figure 1, the empirical distribution of f_2 is skewed to the left. On the other hand, Figure 2 gives the histogram of the empirical distribution of f_2 with the same combination as Figure 1 except for $\underline{\mu}_D = 3\underline{1}$. As displayed in Figure 2, the empirical distribution of f_2 is skewed with a long-tail to the right. When $\underline{\mu}_D = \underline{0}$, W is a linear combination of a central chi-square variable with 1 degree of freedom and a central chi-square variable with $(n-1)$ degrees of freedom. Since W and f_2 have negative correlation, the empirical distribution of f_2 is skewed to the left. If $\underline{\mu}_D = 3\underline{1}$, then $\tau_2 = 0$. Therefore, W is a linear combination of a noncentral chi-square variable with 1 degree of freedom and noncentrality parameter $\tau_1 = 9nI/2(\sigma_e^2 + n\sigma_s^2)$ and a central chi-square

variable with $(n-1)$ degrees of freedom. Comparing with the central chi-square, the density function of a noncentral chi-square variable is shifted to the right by the noncentrality parameter τ_1.

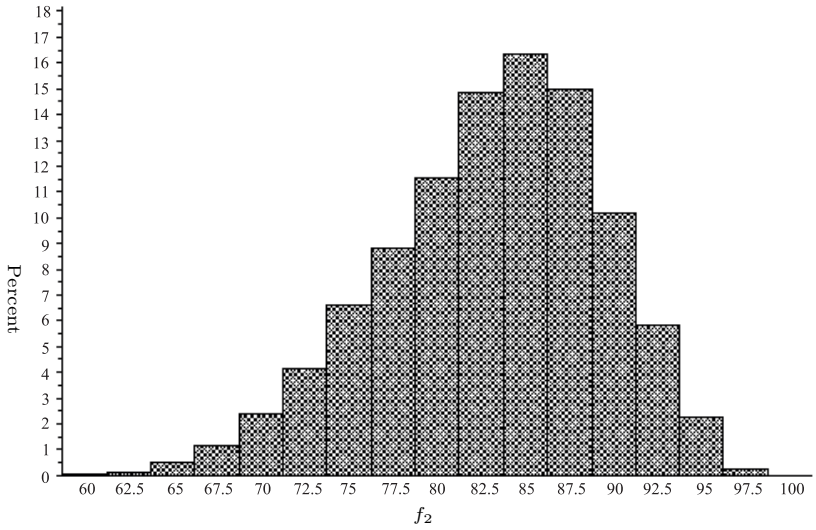

Figure 1. Histogram of f_2. Number of tablets=12 per formulation, number of time points=7, total variability=27.6312, sigma(e)=sigma(s), $\mu = 0$.

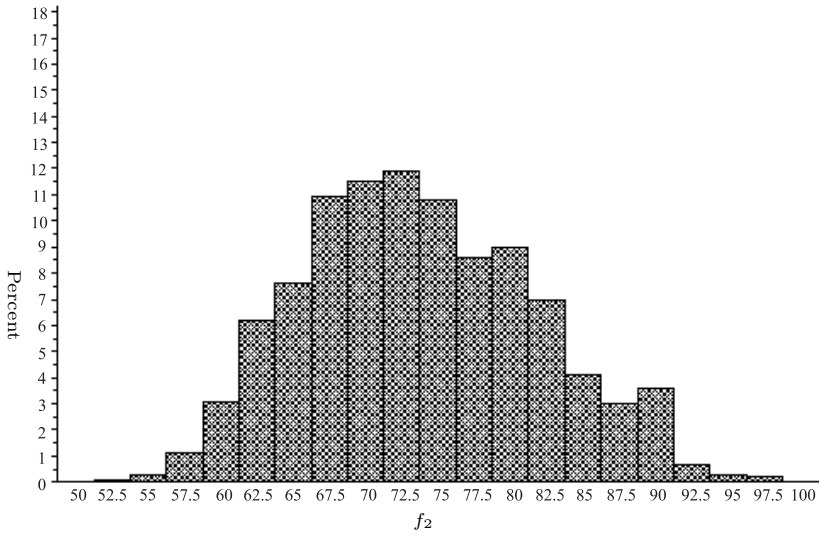

Figure 2. Histogram of f_2. Number of tablets=12 per formulation, number of time points=7, total variability=27.6312, sigma(e)=sigma(s), $\mu = 3$.

Table 1. The mean (the approximate mean (2.6)) and varance (the first term of approximate variance (2.7)) of the empirical distribution of f_2. Total variance = 27.6312, $\underline{\mu}_D = \underline{0}$.

n	I	$K = 0.1$		$K = 0.3$		$K = 0.5$		$K = 0.7$		$K = 0.9$	
		Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
3	6	77.7	67.2	78.1	70.9	78.3	78.5	79.0	92.7	80.5	120.7
		(77.8)	(65.2)	(78.2)	(75.5)	(79.2)	(95.9)	(80.6)	(126.6)	(82.5)	(167.5)
	12	83.7	50.1	83.7	51.9	84.1	57.8	84.6	69.7	85.8	86.3
		(83.8)	(54.1)	(84.2)	(62.6)	(84.9)	(79.6)	(86.1)	(105.0)	(87.7)	(139.0)
	18	86.5	39.2	86.8	41.5	87.1	46.2	87.4	55.6	88.4	69.6
		(86.9)	(45.6)	(87.2)	(52.8)	(87.8)	(67.1)	(88.8)	(88.5)	(90.2)	(117.2)
	24	88.6	32.0	88.7	34.9	89.0	37.8	89.3	45.3	90.0	57.2
		(88.8)	(39.0)	(89.1)	(45.1)	(89.7)	(57.3)	(90.5)	(75.6)	(91.6)	(100.1)
5	6	76.6	40.9	76.9	47.2	77.5	59.0	78.6	78.0	80.2	114.0
		(76.6)	(39.9)	(77.2)	(52.2)	(78.3)	(76.7)	(80.0)	(113.6)	(82.3)	(162.7)
	12	82.7	32.2	82.9	36.6	83.5	44.6	84.3	59.6	85.6	81.6
		(82.8)	(33.1)	(83.3)	(43.3)	(84.2)	(63.7)	(85.6)	(94.2)	(87.5)	(135.0)
	18	85.9	25.0	86.1	29.5	86.6	36.6	87.1	47.3	88.2	66.6
		(86.0)	(27.9)	(86.4)	(36.5)	(87.2)	(53.7)	(88.4)	(79.4)	(90.0)	(113.8)
	24	88.0	21.5	88.2	23.9	88.6	30.0	89.2	40.0	89.9	55.0
		(88.1)	(23.8)	(88.5)	(31.2)	(89.1)	(45.8)	(90.2)	(67.8)	(91.5)	(97.2)
7	6	76.0	29.2	76.5	36.1	76.9	49.7	78.3	70.8	80.2	110.1
		(76.1)	(29.1)	(76.7)	(42.2)	(77.9)	(68.5)	(79.7)	(108.0)	(82.2)	(160.6)
	12	82.4	23.3	82.7	28.6	83.2	38.8	84.2	53.8	85.4	83.5
		(82.4)	(24.1)	(82.9)	(35.0)	(83.9)	(56.8)	(85.4)	(89.6)	(87.4)	(133.2)
	18	85.6	19.3	85.9	23.7	86.3	32.2	87.0	43.1	88.0	65.7
		(85.7)	(20.3)	(86.1)	(29.5)	(87.0)	(47.9)	(88.2)	(75.5)	(89.9)	(112.3)
	24	87.7	15.9	88.0	19.6	88.4	26.8	89.1	36.5	89.9	54.2
		(87.8)	(17.4)	(88.2)	(25.2)	(88.9)	(40.9)	(90.0)	(64.5)	(91.4)	(95.9)

n: Time points
I: The number of tablets
K: The ratio of $\sigma_s^2/(\sigma_s^2 + \sigma_e^2)$

(2) In general, the empirical mean decreases as the total variance increases. The mean of the empirical distribution increases as the ratio of inter-unit variability to the total variability increases or the number of dosage units increases, as shown in Table 1 and Table 2. The empirical means for $\underline{\mu}_D = \underline{0}$ are larger than for $\underline{\mu}_D = 3\underline{1}$. For $\underline{\mu}_D = \underline{0}$, the empirical means obtained by simulation are numerically larger than the approximate mean calculated from (2.6), although they are quite close.

(3) In general, the variance increases as the total variability increases, and decreases as the sample size increases. In addition, it increases as the ratio

of inter-unit variability to the total variability increases or as the number of sampling time points decreases for any fixed combination of sample size and total variability. Large differences between the empirical variance and the first term of (2.7) reflect the fact that the Taylor's expansion of the logarithm diverges if W/n is greater than 1. The empirical variances are larger than the approximate variances when $\underline{\mu}_D = 3\underline{1}$ while they are in general smaller when $\underline{\mu}_D = \underline{0}$.

Table 2. The mean (the approximate mean (2.6)) and variance (the first term of approximate variance (2.7)) of the empirical distribution of f_2. Total variance = 27.6312, $\underline{\mu}_D = 3 \cdot \underline{1}$.

n	I	$K = 0.1$		$K = 0.3$		$K = 0.5$		$K = 0.7$		$K = 0.9$	
		Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
3	6	71.0	74.1	71.6	88.9	72.2	104.4	73.1	124.7	74.3	160.6
		(70.7)	(60.8)	(71.5)	(77.8)	(72.4)	(97.7)	(73.5)	(120.5)	(74.6)	(146.2)
	12	73.1	53.9	73.8	70.5	74.4	84.5	75.1	103.3	75.8	129.9
		(72.9)	(44.6)	(73.6)	(58.1)	(74.2)	(72.8)	(75.0)	(88.8)	(75.8)	(106.0)
	18	73.8	41.5	74.3	55.7	75.1	72.1	75.6	84.6	76.1	104.3
		(73.7)	(34.9)	(74.2)	(45.8)	(74.7)	(57.4)	(75.3)	(69.6)	(75.9)	(82.5)
	24	74.2	34.0	74.7	45.4	75.0	57.6	75.8	74.3	76.1	87.4
		(74.1)	(28.6)	(74.5)	(37.7)	(74.9)	(47.2)	(75.4)	(57.1)	(75.9)	(67.5)
5	6	69.8	47.3	70.8	64.9	71.5	85.1	72.5	109.1	74.1	153.0
		(69.8)	(40.9)	(70.7)	(61.3)	(71.8)	(85.2)	(73.1)	(112.6)	(74.5)	(143.4)
	12	72.3	35.6	73.0	52.7	73.9	70.9	74.8	95.8	75.9	126.8
		(72.3)	(30.5)	(73.0)	(46.7)	(73.9)	(64.3)	(74.7)	(83.5)	(75.7)	(104.2)
	18	73.2	27.3	73.7	43.1	74.5	60.0	75.2	78.6	76.2	103.7
		(73.2)	(24.1)	(73.8)	(37.1)	(74.4)	(51.0)	(75.1)	(65.7)	(75.8)	(81.2)
	24	73.6	21.6	74.2	34.7	74.9	50.0	75.5	66.3	76.1	85.0
		(73.7)	(19.8)	(74.2)	(30.7)	(74.7)	(42.0)	(75.2)	(54.0)	(75.8)	(66.4)
7	6	69.4	35.2	70.3	54.6	71.2	75.7	72.5	101.9	74.0	147.7
		(69.4)	(32.4)	(70.4)	(54.3)	(71.6)	(79.9)	(72.9)	(109.2)	(74.5)	(142.2)
	12	72.1	27.5	73.0	46.3	73.8	65.5	74.6	90.5	75.8	121.9
		(72.0)	(24.5)	(72.8)	(41.8)	(73.7)	(60.7)	(74.6)	(81.3)	(75.6)	(103.4)
	18	73.0	21.1	73.8	38.1	74.4	55.3	75.2	77.0	75.9	100.6
		(73.0)	(19.4)	(73.6)	(33.4)	(74.3)	(48.2)	(75.0)	(64.0)	(75.8)	(80.6)
	24	73.5	17.3	74.1	31.0	74.7	46.6	75.4	63.6	76.3	85.1
		(73.5)	(16.0)	(74.0)	(27.6)	(74.6)	(39.8)	(75.2)	(52.6)	(75.8)	(65.9)

n: Time points
I: The number of tablets
K: The ratio of $\sigma_s^2/(\sigma_s^2 + \sigma_e^2)$

(4) When the total variance is 27.6312 and the average cumulative percent dissolved is between 80 and 90, the coefficient of variation (CV) is between

5.8% and 6.6%. This range of CV represents an upper limit of within-batch variability for most of the immediate released drug products. The empirical probability of concluding similarity is almost 100% for all combinations for $\underline{\mu}_D = \underline{0}$ and $\underline{\mu}_D = 3\underline{1}$. For the total variability of 552.624 which represents a within-batch CV about 30%, the empirical probability of concluding similarity can go above 85% for some combinations for $\underline{\mu}_D = 3\underline{1}$, as shown in Table 3. The empirical probability of concluding similarity increases as the the number of the dosage units increases and decreases as the total variance increases. In addition, it increases as K increases when the total variance is 552.624. The empirical probability of concluding similarity for $\underline{\mu}_D = \underline{0}$ is larger than that for $\underline{\mu}_D = 3\underline{1}$.

Table 3. The empirical probability of concluding similarity between dissolution profiles. Total variance = 552.624, $\underline{\mu}_D = 0$ and $\underline{\mu}_D = \underline{3}$.

n	I	$K = 0.1$		$K = 0.3$		$K = 0.5$		$K = 0.7$		$K = 0.9$	
		$\underline{\mu}_D = \underline{0}$	$\underline{\mu}_D = \underline{3}$	$\underline{\mu}_D = \underline{0}$	$\underline{\mu}_D = \underline{3}$	$\underline{\mu}_D = \underline{0}$	$\underline{\mu}_D = \underline{3}$	$\underline{\mu}_D = \underline{0}$	$\underline{\mu}_D = \underline{3}$	$\underline{\mu}_D = \underline{0}$	$\underline{\mu}_D = \underline{3}$
3	6	34.4	32.8	36.7	34.3	40.1	38.8	46.4	44.9	51.6	51.3
	12	64.5	60.2	64.8	61.8	67.7	64.0	69.8	66.4	69.5	67.9
	18	82.0	76.6	82.4	77.5	81.7	77.7	81.6	77.3	80.3	75.9
	24	90.2	85.3	90.8	85.7	89.1	85.0	88.3	83.3	86.3	82.5
5	6	24.9	24.2	29.0	28.3	35.7	33.7	43.6	43.2	51.9	50.0
	12	64.2	57.1	65.3	60.3	68.9	64.3	70.6	66.6	68.8	66.9
	18	84.3	78.2	84.3	77.8	83.8	78.3	81.7	77.9	80.6	77.2
	24	94.0	88.9	92.8	87.0	90.1	85.8	88.3	84.1	86.6	82.5
7	6	19.9	19.0	24.6	23.1	32.6	30.6	44.3	42.7	51.0	50.8
	12	63.0	56.6	66.1	60.0	69.4	64.4	70.2	68.0	70.6	67.2
	18	86.8	80.1	85.5	79.6	84.3	79.5	82.7	77.4	79.7	77.0
	24	96.1	90.3	94.2	88.0	91.1	86.3	89.3	84.4	85.6	82.8

n: Time points
I: The number of tablets
K: The ratio of $\sigma_s^2/(\sigma_s^2 + \sigma_e^2)$

4. Mean Squared Distance and Mean Absolute Difference

In general, the metrics for assessing dissolution similarity can be classified into two groups. The first group is based on functions of absolute differences in population averages of cumulative percent dissolved between test and reference formulations. The other is based on functions of the squared differences in population averages between test and reference formulations. That is, $g(\sum_{k=1}^{n} w_k(\mu_{Tk} - \mu_{Rk})^2)$ or $g(\sum_{k=1}^{n} w_k|\mu_{Tk} - \mu_{Rk}|)$, where w_k is the weight of

the kth time point. The sum of weights over all time points should be 1. For example if $w_k = 1/n$, then they are functions of mean squared distance and mean absolute difference, respectively. We only consider the simplest case of equal weights and identity function. Note that θ in (2.8) is also based on mean squared distance. We shall see that even for the simplest case, the inference based on $\sum_{k=1}^{n}(\mu_{Tk} - \mu_{Rk})^2$ or $\sum_{k=1}^{n}|\mu_{Tk} - \mu_{Rk}|$ can be very complicated. The proofs of the following theorems and lemmas are straightforward and can be obtained from the authors upon request.

Theorem 1. *Under the assumption of multivariate normality with compound symmetry covariance, we have the following properties:*

(1) *Let* $SS_W = 2[\sum_{i=1}^{I}(X_{Tki} - \bar{X}_{Tk.})^2 + \sum_{i=1}^{I}(X_{Rki} - \bar{X}_{Rk.})^2]$. *An unbiased estimator of* $\sum_{k=1}^{n}(\mu_{Tk} - \mu_{Rk})^2$ *is given by*

$$T = \sum_{k=1}^{n}[(\bar{X}_{Tk.} - \bar{X}_{Rk.})^2 - \frac{1}{2I(I-1)}SS_W]; \tag{4.1}$$

(2) *the distribution of* T *is that of a linear combination of independent noncentral chi-square random variables:* $\frac{2}{I}[(\sigma_e^2 + n\sigma_s^2)\chi^2(1,\tau_1) + \sigma_e^2\chi^2(n-1,\tau_2)] - \frac{1}{I(I-1)}\{\sigma_e^2\chi^2[2(n-1)(I-1)] + (\sigma_e^2 + n\sigma_s^2)\chi^2[2(I-1)]\}$.

If the targeted parameter of f_2 based on the squared differences in population averages is θ, then an estimate of θ is $\hat{\theta} = 100 - 25\log_{10}[1 + \frac{1}{n}T]$. The fact that T can be negative makes its distribution very complicated.

An alternative test statistic may be given as

$$T_a = \frac{\sum_{k=1}^{n}(\bar{X}_{Tk.} - \bar{X}_{Rk.})^2}{\frac{1}{I(I-1)}\sum_{k=1}^{n}[\sum_{i=1}^{I}(X_{Tki} - \bar{X}_{Tk.})^2 + \sum_{i=1}^{I}(X_{Rki} - \bar{X}_{Rk.})^2]} = \frac{W}{SS_W/2I(I-1)}. \tag{4.2}$$

The numerator W was defined earlier and the denominator is the sum of the pooled within group sums of squares over all time points. Under the assumption of normality with constant difference for all time points and a compound symmetry covariance, the distribution of T_a is that of a ratio of two linear combinations of noncentral chi-square variables:

$$\frac{\frac{2}{I}(\sigma_e^2 + n\sigma_s^2)\chi^2(1,\tau_1) + \frac{2}{I}\sigma_e^2\chi^2(n-1,\tau_2)}{\frac{1}{I(I-1)}\{\sigma_e^2\chi^2[2(n-1)(I-1)] + (\sigma_e^2 + n\sigma_s^2)\chi^2[2(I-1)]\}}. \tag{4.3}$$

Let

$$\theta_a = \frac{E(W)}{E(SS_W)/2I(I-1)} = \frac{I\sum_{k=1}^{n}\mu_{Dk}^2}{2n(\sigma_e^2 + \sigma_s^2)} + 1. \tag{4.4}$$

From equation (4.4), $\theta_a \geq 1$ and $\theta_a = 1$ only if $\underline{\mu}_D = 0$. If $\sigma_s^2 = 0$ then T_a follows a noncentral F distribution with degrees of freedom n and $2n(I-1)$ and noncentrality parameter $\tau_1 + \tau_2$. Hence, θ_a is reduced to $(n + \tau_1 + \tau_2)/n$ and $E(T_a) = \frac{(I-1)(n+\tau_1+\tau_2)}{[n(I-1)-1]}$. Then $\frac{[n(I-1)-1]}{[n(I-1)]}T_a$ is an unbiased estimator of θ_a. If $\sigma_e^2 = 0$ then T_a follows a noncentral F distribution with degrees of freedom 1 and $2(I-1)$ and noncentrality parameter τ_1. Hence, θ_a is reduced to $1 + \tau_1$ and $E(T_a) = \frac{(I-1)(1+\tau_1)}{(I-2)}$. Then $\frac{(I-2)}{(I-1)}T_a$ is an unbiased estimator of θ_a. If $\sigma_s^2 \neq 0$ and $\sigma_e^2 \neq 0$ then the evaluation of $E(T_a)$ becomes very difficult. The assessment of similarity between dissolution profiles based on θ_a can again be formulated as testing a one-sided hypothesis:

$$H_0 : \theta_a \geq \Delta \qquad v.s. \qquad H_a : \theta_a < \Delta, \tag{4.5}$$

where Δ is the prespecified upper similarity limit determined by cumulative average difference and both intra and inter-unit variabilities. Further investigation of T_a is needed.

Next, we investigate inference based on the metric $\sum_{k=1}^{n} |\mu_{Tk} - \mu_{Rk}|$. We start with a single sampling time point, i.e., $n = 1$.

Lemma 1. *If $\bar{X}$ follows a normal distribution with mean μ and variance σ^2/I, then*

(1) *$|\bar{X}|$ is asymptotically unbiased for $|\mu|$ as $I \to \infty$;*

(2) *$|\bar{X}|$ is a consistent estimator of $|\mu|$ for $\mu \neq 0$;*

(3) *The moment generating function of $|\bar{X}|$ is given as*

$$e^{\mu t + \frac{\sigma^2 t^2}{2I}} \Phi\left(\frac{t\sigma}{\sqrt{I}} + \frac{\mu\sqrt{I}}{\sigma}\right) + e^{-\mu t + \frac{\sigma^2 t^2}{2I}} \Phi\left(\frac{t\sigma}{\sqrt{I}} - \frac{\mu\sqrt{I}}{\sigma}\right),$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution;

(4) *An unbiased estimator of $|\mu|$ does not exist (the proof is given in Appendix);*

(5) *The asymptotic distribution of $|\bar{X}|$ is, for large I, $|\bar{X}| \sim N(|\mu|, \sigma^2/I)$ for $\mu \neq 0$;*

(6) *The bias of $|\bar{X}|$ is given by*

$$E(|\bar{X}|) - |\mu| = \begin{cases} \sqrt{\frac{2}{\pi}}\frac{\sigma}{\sqrt{I}}, & \text{if } \mu = 0, \\ 2\mu[\Phi(\frac{\sqrt{I}\mu}{\sigma}) - 1] + 2\frac{\sigma}{\sqrt{I}}\phi(\frac{\sqrt{I}\mu}{\sigma}), & \text{if } \mu > 0, \\ 2\mu\Phi(\frac{\sqrt{I}\mu}{\sigma}) + 2\frac{\sigma}{\sqrt{I}}\phi(\frac{\sqrt{I}\mu}{\sigma}), & \text{if } \mu < 0, \end{cases} \tag{4.6}$$

where $\phi(\cdot)$ denotes the probability density function of the standard normal distribution;

(7) *The variance of* $|\bar{X}|$ *is given as* $\mu^2+\sigma^2/I-\{\mu[2\Phi(\frac{\sqrt{I}\mu}{\sigma})-1]+2\frac{\sigma}{\sqrt{I}}\phi(\frac{\sqrt{I}\mu}{\sigma})\}^2$.

Let $Z=\sqrt{I}(|\bar{X}|-|\mu|)/\sigma$ be the standardized form of $|\bar{X}|$. Then the moment generating function of Z is

$$M_Z(t)=e^{-t\sqrt{I}|\mu|/\sigma}M_{|\bar{X}|}(\sqrt{I}t/\sigma)\longrightarrow\begin{cases}e^{\frac{1}{2}t^2}, & \text{if } \mu>0,\\ 2e^{\frac{1}{2}t^2}\Phi(t), & \text{if } \mu=0,\\ e^{\frac{1}{2}t^2}, & \text{if } \mu<0, \text{ as } I\to\infty.\end{cases}$$

However, when $\mu=0$, the asymptotic distribution of the standardized form Z does not converge to the standardized normal distribution as $I\to\infty$.

Assume that the $\bar{X}_{Tk.}-\bar{X}_{Rk.}$, $k=1,\ldots,n$, follow a multivariate normal distribution with mean $\mu_{Dk}=\mu_{Tk}-\mu_{Rk}$, $k=1,\ldots,n$, and compound symmetry covariance Σ. Let σ_{kk} be the kth diagonal element of Σ. By lemma 1, we have the following properties:

(1) $T_r=\sum_{k=1}^n|\bar{X}_{Tk.}-\bar{X}_{Rk.}|$ is a consistent estimator of $\sum_{k=1}^n|\mu_{Tk}-\mu_{Rk}|$;
(2) The bias of T_r is negligible if the number of individual dosage units is large, that is, T_r is asymptotically unbiased;
(3) An unbiased estimator of $\sum_{k=1}^n|\mu_{Tk}-\mu_{Rk}|$ does not exist;
(4) The moment generating function of $|D_k|=|\bar{X}_{Tk.}-\bar{X}_{Rk.}|$ is

$$e^{\mu_{Dk}t+\frac{1}{2}\sigma_{kk}^2t^2}\Phi(t\sigma_{kk}+\frac{\mu_{Dk}}{\sigma_{kk}})+e^{-\mu_{Dk}t+\frac{1}{2}\sigma_{kk}^2t^2}\Phi(t\sigma_{kk}-\frac{\mu_{Dk}}{\sigma_{kk}}),\quad for\quad k=1,\ldots,n;$$

(5) The expectation of T_r is $\sum_{k=1}^n\{\mu_{Dk}[2\Phi(\frac{\mu_{Dk}}{\sigma_{kk}})-1]+2\sigma_{kk}\phi(\frac{\mu_{Dk}}{\sigma_{kk}})\}$.

If $\underline{\mu}_D=\underline{0}$, for large I, the limit of moment generating function of the standardized form $Z_k=(|D_k|-|\mu_{Dk}|)/\sigma_{kk}$ is given by $M_{Z_k}(t)=2e^{\frac{1}{2}t^2}\Phi(t)$, $E(Z_k)=M'_{Z_k}(0)=\sqrt{\frac{2}{\pi}}$ and $\mathrm{Var}\,(Z_k)=M''_{Z_k}(0)-(E(Z_k)^2)=1-\frac{2}{\pi}$. On the other hand, the moment generating function of T_r is complicated because $|\bar{X}_{Tk.}-\bar{X}_{Rk.}|$, $k=1,\ldots,n$, are correlated.

If $\mu_{Dk}\neq 0$ for all time points, use the delta-method with $g(y)=|y|$ to get $|\underline{D}|=(|\bar{X}_{T1.}-\bar{X}_{R1.}|,\ldots,|\bar{X}_{Tn.}-\bar{X}_{Rn.}|)'\sim N(|\underline{\mu}_D|,\Sigma^*)$, where $|\underline{\mu}_D|=|\underline{\mu}_T-\underline{\mu}_R|$, $\Sigma^*=(\frac{\partial|D_k|}{\partial D_k}|_{\underline{D}=\underline{\mu}_D})'\Sigma(\frac{\partial|D_k|}{\partial D_k}|_{\underline{D}=\underline{\mu}_D})$ and

$$(\frac{\partial|D_k|}{\partial D_k}|_{\underline{D}=\underline{\mu}_D})=\begin{cases}1, & \text{if } \mu_{Dk}>0,\\ -1, & \text{if } \mu_{Dk}<0.\end{cases}$$

Since $T_r=\sum_{k=1}^n|\bar{X}_{Tk.}-\bar{X}_{Rk.}|=\underline{1}'|\underline{D}|$, the asymptotic distribution of T_r is then given as $N(\sum_{k=1}^n|\mu_{Tk}-\mu_{Rk}|,\underline{1}'\Sigma^*\underline{1})$, where $\underline{1}'\Sigma^*\underline{1}=\sum_{k=1}^n\sigma_{kk}+2\sum\sum_{i<j}\sigma_{ij}$ $(\frac{\partial|D_i|}{\partial D_i}|_{\underline{D}=\underline{\mu}_D})'$ $(\frac{\partial|D_j|}{\partial D_j}|_{\underline{D}=\underline{\mu}_D})$. So the asymptotic variance of T_r depends on not only the sign of $\underline{\mu}_D$ but also on whether $\mu_{Dk}=0$, $k=1,\ldots,n$.

5. Discussion

The *SUPAC-IR* issued on November 30, 1995 by the U.S. FDA adopted the similarity factor f_2 proposed by Moore and Flanner (1996) as one criterion for assessing dissolution similarity. Any criteria for assessing similarity between dissolution profiles should be based on population parameters. However, f_2 is a sample statistic obtained from dissolution testing, and it will change with the observed time. In addition, the corresponding population parameter that f_2 tries to estimate is unknown, and the consumer's and the producer's risk is very difficult to evaluate. On the other hand, if one takes f_2 as a biased estimate for θ then the consumer's and the producer's risk may be evaluated. If one takes f_2 as an estimate for equation (2.6) then the criterion based on f_2 is in fact a one-sided hypothesis. If one takes f_2 as an estimate of θ then it tests two-sided similarity.

The distribution of f_2 is unnecessarily complicated in order to produce an acceptable range of 50 to 100. As a result its sampling error can not be analytically quantified, and its expectation and variance require numerical integration. In addition, a value of f_2 between 50 and 100 computed from a sample of 24 dosage units does not guarantee dissolution similarity in a population of millions of dosage units because of sampling error. Thus the similarity factor is based on convenience, although it does measure the similarity of two dissolution profiles and is easy to perform. The proposed critical value 50 was based on reviewers' practical understanding of similarity of two true profiles. However, f_2 as an estimate does not have the desirable properties often expected in statistical science.

If the targeted parameter is based on functions of the mean squared distance in population averages of cumulative percent dissolved, unbiased estimators of the mean squared distance in population averages of cumulative percent dissolved can be derived as shown in Section 4. Its distribution, however, is also quite complicated and its observed value may be negative. On the other hand, one can directly apply W/n or Mahalanobis distance to assess similarity of dissolution profiles. Mean squared distance W/n is easier to compute than Mahalanobis distance. But it needs an estimate of the covariance for the sampling distribution. Hence it is much more complicated for computing the sampling distribution and confidence limits. Mahalanobis distance is, on the other hand, the standardized form of W/n with a well-known sampling distribution. It involves the estimation of the covariance matrix in calculating the point estimate but not the sampling distribution. Hence, the small sample problem for covariance matrix estimation in Mahalanobis distance would be exactly the same as for W/n. Asymptotically,

the sample Mahalanobis distance is unbiased for the corresponding population parameter (Anderson (1984)). Since the sampling distribution of W/n is well understood, statistical inference based on W/n under (2.9) can be easily carried out. However, W/n is not scaled and is not invariant to scale change. In addition, the sampling distribution of W/n involves estimation of the noncentrality parameters of noncentral chi-square distributions. As demonstrated in the previous section, the sampling distributions of estimators for criteria based on the mean squared distance are quite complicated. Bootstrap techniques may therefore be an attractive approach to evaluating the sampling distributions of $\hat{\theta}$, T_a and T_r.

Acknowledgements

We would like to thank two referees and the Associate Editor for their constructive suggestions and insightful comments which greatly improved the paper.

Appendix

Proof of equation (2.3). Since $\underline{D} \sim N(\underline{\mu}_D, \Sigma)$ and $W = \underline{D}'\underline{D}$, by properties of multivariate normality, W can be written as a linear combination of independent noncentral chi-square random variables: $W = \sum_{i=1}^{n} \lambda_i Y_i$, where $\lambda_i^{'}$s are the eigenvalues of Σ and Y_i follows a noncentral chi-square distribution with 1 degree of freedom and noncentral parameter $\tau_i = \left(\underline{\mu}_D' \underline{e}_i\right)^2 / \lambda_i$, and $\underline{e}_i$ is the eigenvector of Σ corresponding to λ_i.

Because the joint density function of $Y_1, \ldots, Y_n$ can be written as

$$f(y_1, \ldots, y_n) = \prod_{i=1}^{n} f_i(y_i) = \prod_{i=1}^{n} \left[\sum_{k=0}^{\infty} g_k(\tau_i) f_{1+2k}(y_i) \right],$$

where $g_k(\tau_i) = e^{-\tau_i/2} \times \frac{(\tau_i/2)^k}{k!}$, and $f_{1+2k}(y_i)$ is the density function of the central chi-square distribution with $1 + 2k$ degrees of freedom. Then if $S_j = Y_j$ for $j = 1, \ldots, n-1$, and $W = \sum_{i=1}^{n} \lambda_i Y_i$, it follows that the joint density function of $(S_1, \ldots, S_{n-1}, W)$ is

$$f(s_1, \ldots, s_{n-1}, w) = \prod_{i=1}^{n-1} \left[\sum_{k=0}^{\infty} g_k(\tau_i) f_{1+2k}(s_i) \right] \cdot \left[\sum_{k=0}^{\infty} g_k(\tau_n) f_{1+2k} \left(\frac{1}{\lambda_n} \left(w - \sum_{i=1}^{n-1} \lambda_i s_i \right) \right) \right] / \lambda_n.$$

Then (2.3) follows from $f_2 = 100 - 25 \times \log_{10}(1 + \frac{W}{n})$.

Proof of Lemma 1. (4) If Y was a normal distribution with mean μ and variance σ^2, there is no unbiased estimator of $|\mu|$.

Proof. Suppose there exists an unbiased estimator of $|\mu|$, say $g(Y)$. Because $g(Y)$ is an unbiased estimator, it must be a function of Y and is not a function of μ such that $E(g(Y)] = |\mu|$. If $\mu \geq 0$ then $E[g(Y)] = \mu$ and $E(Y) = \mu$. If $\mu \geq 0$ then $E[g(Y)] = \mu$ and $E(Y) = \mu$. Since Y is complete sufficient for $\mu \geq 0$, $E[g(Y) - Y] = 0$ implies $g(Y) = Y$ almost surely for $\mu \geq 0$. Similarly $g(Y) = -Y$ almost surely for $\mu < 0$. Then $g(Y) = YI_{(\mu \geq 0)} - YI_{(\mu < 0)}$ and $g(Y)$ depends on μ, a contradiction.

References

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis.* 2nd edition. John Wiley, New York.

Chow, S. C. and Ki, F. Y. C. (1997). Statistical comparison between dissolution profiles of drug product. *J. Biopharm. Statist.* **7**, 241-258.

Gill, J. L. (1988). Repeated measurement: split-plot trend analysis versus analysis of first difference. *Biometrics* **44**, 289-297.

Guidance for Industry: Immediate Release Solid Oral Dosage Forms; Scale-up and Postapproval Changes: Chemistry, Manufacturing, and Controls; In Vitro Dissolution Testing; In Vivo Bioequivalence Documentation. Center for Drug Evaluation and Research, Food and Drug Administration, November 1995.

FDA Guidance for Industry (1997): *Dissolution Testing of Immediate Release Solid Oral Dosage Forms. Dissolution Technology* **4**, 15-22.

FDA Guidance for Industry (1997) *Extended Release Solid Oral Dosage Forms; Development, Evaluation and Application of In Vitro/In Vivo Correlation. Dissolution Technology* **4**, 23-32.

FDA Guidance for Industry (1997): *Modified Release Solid Oral Dosage Forms; Scale-up and Postapproval Changes (SUPAC-MR); Chemistry, Manufacturing and Controls; In Vitro Dissolution Testing and In Vivo Bioequivalence Documentation.* U.S. FDA, Rockville, Maryland.

Johnson, N. L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions.* Vol. 2. John Wiley, New York.

Lesson, L. J. (1995). In *vitro/vivo* correlation. *Drug Information J.* **29**, 903-915.

Liu, J. P., Ma, M. C. and Chow, S. C. (1997). Statistical evaluation of similarity factor f_2 as a criterion for assessment of similarity between dissolution profiles. *Drug Information J.* **31**, 1255-1271.

Moore, J. W. and Flanner, H. H. (1996). Mathematical comparison of curves with an emphasis on dissolution profiles. *Pharmaceutical Technology* **20**, 64-74.

O'Hara, T., Dunne, A., Kinahan, A., Cunningham, S., Stark, P. and Devane, J. (1996). Review of methodologies for the comparison of dissolution profile data. In *In Vitro-In vivo Correlation* (Edited by D. Young, J. D. Devane and J. Butler). Plenum Publishing, New York.

Polli, J. E., Rekhi, G. S., Augsburger, L. L. and Shah, V. P. (1997). Methods to compare dissolution profiles and a rationale for wide dissolution specifications for metoprolol tartate tablets. *J. Pharm. Sci.* **86**, 690-700.

Sathe, P., Tsong, Y. and Shah, V. P. (1996a). *In vitro* dissolution profile comparison: statistics and analysis, model dependent approach. *Pharmaceutical Research* **13**, 1799-1803.

Sathe, P., Tsong, Y. and Shah, V. P. (1996b). *In vitro* dissolution profile comparison and IVIVR: Carbamazepine case. In *In Vitro-In vivo Correlation* (Edited by D. Young, J. D. Devane and J. Butler). Plenum Publishing, New York.

Tsong, Y., Hammerstrom, T., Sathe, P. and Shah, V. P. (1996). Statistical assessment of mean differences between two dissolution data sets. *Drug Information J.* **30**, 1105-1112.

Tsong, Y., Hammerstrom, T., Sathe, P. and Shah, V. P. (1997). Comparing two dissolution data sets for similarity. *Amer. Statist. Assoc.*, 1996 Proceedings of the Biopharmaceutical Section, 129-134.

Department of Statistics, National Cheng-kung University, Tainan 701, Taiwan.

E-mail: mcma@ibm.stat.ncku.edu.tw

Division of Biostatistics, National Health Research Institutes, Taipei 115, Taiwan.

E-mail: jpliu@nhri.org.tw

(Received September 1997; accepted December 1998)