Taylor & Francis
Taylor & Francis Group

A MULTIVARIATE TEST FOR SIMILARITY OF TWO DISSOLUTION PROFILES

H. Saranadasa

Ortho McNeil Pharmaceutical, Inc., Raritan, New Jersey, USA

K. Krishnamoorthy

University of Louisiana at Lafayette, Lafayette, Louisiana, USA

A multivariate test of size α for assessing the similarity of two dissolution profiles is proposed. The inferential procedure is developed by using the approach for the common mean problem in a multivariate setup due to Halperin (1961). The performance of the proposed method is compared with Intersection Union Test as well as f_2 criterion recommended by the FDA through a simulation study. All the methods are illustrated with real examples.

1. INTRODUCTION

Dissolution testing is performed on 6 or 12 dosage units (assume tablets or capsules are the dosage form) and placing them in agitated media. The dissolution test system has six vessels, each holding a liter of media. A rotating basket or paddle is lowered to agitate the contents, after which a tablet is dropped into the vessel and dissolution samples are collected at different time points (e.g., every 15 minutes in the first hour and every 60 minutes after the first hour) and analyzed, usually via chromatography or UV spectroscopy. The response at time t is the cumulative amount (%) of drug released into the media. The dissolution profiles for solid dosage forms are developed in connection with observations taken on tablets or capsules over time. It is the curve of the mean dissolution rate (cumulative % dissolved) over time. The pharmaceutical scientists are interested in a comparision of these profiles under different conditions related to formulation forms, lot-to-lot and brand-to-brand variation. For example, if dissolution profile similarity is demonstrated between the prechange drug product and the postchange formulation, in vivo bioequivalence testing can be waived for most changes.

It is a big advantage for the industry to avoid conducting clinical studies to demonstrate bioequivalency, and these studies are not only time-consuming but also expensive. For some drugs, bioavailability needs to be demonstrated only if the product fails to achieve adequate dissolution compared with a test standard. Also in the manufacturing phase, dissolution profile of a new lot (may be some changes with respect to chemical manufacturing and control) will be tested against the validation lots for similar dissolution profiles to ensure the compliance. Therefore, it is of great interest to the pharmaceutical scientist to compare dissolution profiles. The statistical challenge is how to define and test the two population dissolution profiles that are "similar" based on the sample dissolution data collected over time.

The U.S. Food and Drug Administration (FDA) has issued several guidelines describing circumstances for which scale-up and postapproval change (SUPAC) in the components of drug product manufacturing site or the manufacturing process and equipment of formulation are acceptable. One requirement is to establish the similarity of dissolution across a suitable time interval.

The U.S. FDA's guidance for industry on dissolution testing of immediate-release (IR) solid oral dose forms (1997), as well as SUPAC-IR (1995), SUPAC-MR (1997), and bioavailability and bioequivalence study guidance for oral dosage forms, describes the model independent mathematical approach proposed by Moore and Flanner (1996) for calculating a dissimilarity factor (f_1) and a similarity factor (f_2) of dissolution across a suitable time interval. The similarity factor (f_2) (where $0 \leq f_2 \leq 100$ and $f_2 \geq 50\%$ implies dissolution profiles are similar) is a function of mean differences and does not take into account the differences in dissolution within the test and reference batches. Hence, careful interpretation is warranted when f_2 is used as a similarity factor when the variances of the profiles are very different.

Previous articles have discussed the more serious deficiencies of using the f_2 factor for assessing the similarity between two profiles. One of the major drawbacks identified was finding the sampling distribution of the statistics. This statistic has complicated properties and deriving the distribution of the statistic is not mathematically tractable. Shah et al. (1998) proposed a bootstrap method to calculate a confidence interval for the f_2 factor. Because f_2 is sensitive to the measurements obtained after either the test or reference batch has dissolved more than 85%, Shah et al. (1998) recommended a limit of one sampling time point after 85% dissolution. A recent article, discussing the aspects of the dissolution profile testing problem, by Eaton et al. (2003) also raised some issues concerning the use of the f_2 statistic.

Several other authors discussed criteria for statistical evaluation of similarity of dissolution profiles using model-independent as well as model-dependent approaches. Some of the model-dependent approaches were discussed by Sathe et al. (1996). Tsong et al. (1997) and Chow and Ki (1997) proposed methods based on autoregressive time series models. Wang et al. (1999) showed that the methods using (1) Intersection Union Test (IUT) (2) the Likelihood Ratio Test (LRT), and (3) inversion of the Hotelling's T^2 confidence region are all equivalent. The hypothesis-testing procedures are based on f_2 criterion, and resampling by bootstrap method has been discussed by Shah et al. (1998) and Ma et al. (1999, 2000).

In the FDA guideline for industry, the procedure allows the use of mean data and recommends that the Relative Standard Deviation (RSD) at an earlier time point (for example 5 or 10 minutes) not be more than 20%, and at other time points

not more than 10%. In instances where the RSD within a batch is more than 15%, the guideline suggests using a multivariate model-independent procedure, but no references are given for these situations.

The methods proposed by Tsong et al. (1996) and Saranadasa (2001) are based on a confidence set method using a multivariate normal distribution. In this article, we consider Saranadasa's (2001) proposal and give an exact solution for establishing similarity of dissolution profiles. Saranadasa (2001) approach will be briefly discussed at the end of Section 4. The solution is based on our observation that the present problem, in the setup of Tsong et al. (1996) and Saranadasa (2001), is equivalent to the hypothesis-testing problem for the common mean of a multivariate normal distribution discussed by Halperin (1961) and Krishnamoorthy and Lu (2005).

Let us first introduce some notations to understand the two multivariate tests proposed in the literature and which will be presented in next two sections.

Let u_{ijk} be the observed cumulative percent dissolved for dosage unit j at sampling time t_k for formulation i, where $i = 1, 2$ (1: Reference formulation and 2: Test formulation), $j = 1, 2, \ldots, n_i$, and $k = 1, 2, \ldots, p$ (p time points). Let $\bar{u}_1$ and $\bar{u}_2$ denote the sample mean vectors of length p of the reference profile and test profile respectively. Then $\bar{u}_d = \bar{u}_1 - \bar{u}_2 = (\bar{u}_{d1}, \ldots, \bar{u}_{dp})'$ be the sample mean difference vector and the corresponding population vector is defined as $\mu = \mu_1 - \mu_2$. The problem of interest be to test the hypotheses

$$H_0 : \mu_i > \delta_0 \text{ or } \mu_i < -\delta_0 \text{ for some } i \quad \text{vs.} \quad H_a : -\delta_0 \leq \mu_i \leq \delta_0 \text{ for all } i, \quad (1)$$

where δ_0 is the prespecified acceptable dissolution profile difference and μ_i's are the components of vector μ. To establish the similarity of dissolution across a suitable time interval, the value of $\delta_0 = 10$.

2. THE FIT FACTOR f_2

The FDA promoting fit factor, f_2, is a mathematical index ($0 \leq f_2 \leq 100$) constructed by a function of Euclidean distance of population dissolution mean difference vectors of test and reference formulations. As defined earlier, $\mu = \mu_1 - \mu_2$ (μ_1 and μ_2 are the population mean vectors of length p of reference and test formulation, respectively). The f_2 is defined by Moore and Flanner (1996) as follows:

$$f_2 = 50 \log_{10}\left(100\left(1 + \frac{\mu'\mu}{p}\right)^{-\frac{1}{2}}\right). \quad (2)$$

Notice that $f_2 = 100$ when the two dissolution profiles are identical (i.e., $\mu'\mu = 0$). If the dissolution of one formulation is completed (100%) before the other begins, then $f_2 = 50 \log_{10}\left(100(1 + 100^2)^{\frac{1}{2}}\right) = -0.001 \sim 0$. If $|\mu_{1i} - \mu_{2i}| = 10$ for all $i = 1, 2, \ldots, p$, f_2 is very close to 50 and dissolution profiles with $|\mu_{1i} - \mu_{2i}| \leq 10$ (or $f_2 \geq 50\%$) for all $i = 1, 2, \ldots, p$ are considered as similar dissolution profiles according to the FDA guideline for IR solid oral dosage form (1997).

3. INTERSECTION UNION TEST (IUT)

As we mentioned in the earlier section [see Wang et al., 1999 for more details], the intersection union principal leads to a test of the form: Reject dissimilar dissolution profiles if $Max\{|d_k| + c(\frac{S_k^2}{n})^{\frac{1}{2}}\} < \delta_0$ for all time points t_k, where $|d_k|(|\bar{u}_{dk}|)$ is the observed absolute mean difference at the t_kth time point, δ_0 is the prespecified acceptable dissolution profile difference, S_k^2 is the pooled variance of the two dissolution profiles at the t_kth time point, n is the degrees of freedom $(n_1 + n_2 - 2)$, and c is the $(1-\alpha\%)$ percentile of the Student's t distribution with n degrees of freedom. Berger and Hsu (1996) showed that this is a size α test.

4. THE PROPOSED MULTIVARIATE APPROACH

Suppose $u_{ij} \sim N_p(\mu_i, \Sigma)$, $j = 1, 2, \ldots n_i$ and $i = 1, 2$, are two independent multivariate normal dissolution profiles, one being a reference batch or a prechange batch and the other being a test batch or postchange batch. Samples were taken at p different common time points; n_i values are the number of tablets used in reference and test batches, normally $n_1 = n_2$ and the common sample size is 6 or 12. Let $\bar{u}_1$ and $\bar{u}_2$ denote the sample mean vectors of length p of the reference profile and test profile, respectively; let S_1 and S_2 denote the sample variance-covariance matrices of the reference profile, and test profile respectively. Then

$$\bar{u}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} u_{ij} \quad \text{and} \quad S_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (u_{ij} - \bar{u}_i)(u_{ij} - \bar{u}_i)', \quad i = 1, 2 \qquad (3)$$

To establish similar dissolution profiles, let us assume that $\mu_1 - \mu_2 = e\delta$, where e denotes the $p \times 1$ vector of ones, and δ is an unknown constant. This implies that the sample mean difference vector, $\bar{u}_d = \bar{u}_1 - \bar{u}_2$, of two profiles,

$$\bar{u}_d \sim N_p\left(e\delta, \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\Sigma\right)$$

independently of $V = (n_1 + n_2 - 2)S = (n_1 - 1)S_1 + (n_2 - 1)S_2 \sim W_p(n_1 + n_2 - 2, \Sigma)$, where $W_p(n, \Delta)$ denotes the Wishart distribution with degrees of freedom n and the scale matrix Δ. We here note that S is an unbiased estimator of Σ.

Notice that to establish similar dissolution profiles, it is enough to find sample evidence in favor of H_a: $-\delta_0 \leq \delta \leq \delta_0$. Thus, we want to test

$$H_0: \delta > \delta_0 \text{ or } \delta < -\delta_0 \quad \text{vs.} \quad H_a: -\delta_0 \leq \delta \leq \delta_0 \qquad (4)$$

We like to point out that whenever H_a in Eq. (4) holds, then the H_a in Eq. (1) also holds. The above hypothesis-testing problem is a two-sample version of the one-sample "common mean" problem considered by Halperin (1961) and recently by Krishnamoorthy and Lu (2005). In the one-sample case, we have a multivariate normal population with mean vector $\mu = e\delta$ and unknown covariance matrix Σ, and the problem of interest is to develop inferential procedures about δ based on a sample of observations. Therefore, solutions to the present problem can be easily obtained from the results of the studies just cited.

If Σ is known, then the best linear unbiased estimator for δ is given by $e'\Sigma^{-1}\bar{u}_d/(e'\Sigma^{-1}e)$. If Σ is unknown, then replacing it by the sample covariance matrix S, we get the following natural estimator

$$\hat{\delta} = \frac{e'S^{-1}\bar{u}_d}{e'S^{-1}e} = \frac{e'V^{-1}\bar{u}_d}{e'V^{-1}e} \tag{5}$$

of δ which is also the maximum likelihood estimator. As pointed out by Halperin (1961), the following transformation is useful to describe the distribution of $\hat{\delta}$. Let

$$\bar{y} = \bar{u}_{d1}, \quad \bar{x}_1 = \bar{u}_{d1} - \bar{u}_{d2}, \ldots, \bar{x}_{p-1} = \bar{u}_{d1} - \bar{u}_{dp}$$

where $\bar{u}_{di}$ denotes the ith component of $\bar{u}_d$. In matrix notation, we write this transformation as

$$\left(\frac{\bar{y}}{\bar{X}}\right) = A\bar{u}_d, \quad \text{where } A_{p\times p} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & -1 & 0 & \ldots & 0 \\ 1 & 0 & -1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0 & 0 & \ldots & -1 \end{pmatrix}$$

Write

$$A = \begin{pmatrix} a_{11} & 0 \\ a_{21} & A_{22} \end{pmatrix} \text{ and } V = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & V_{22} \end{pmatrix}$$

so that a_{11} and v_{11} are scalars and

$$AVA' = \begin{pmatrix} v_{11} & a_{11}v_{11}a'_{21} + a_{11}v_{12}A_{22} \\ a_{21}v_{11}a_{11} + A_{22}v_{21}a_{11} & a_{21}v_{11}a'_{21} + A_{22}v_{21}a'_{12} + a_{12}v_{12}A_{22} + A_{22}V_{22}A'_{22} \end{pmatrix}$$

$$= \begin{pmatrix} w_{yy} & w_{yX} \\ w_{Xy} & W_{XX} \end{pmatrix}, \text{ say}$$

In terms of these notations, we can write $\hat{\delta}$ as

$$\hat{\delta} = \bar{y} - w_{yX}W_{XX}^{-1}\bar{X}$$

Let

$$Q = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1}\bar{X}'W_{XX}^{-1}\bar{X}$$

Then, $Q \sim \frac{p-1}{n_1+n_2-p}F_{p-1,n_1+n_2-p}$, where $F_{a,b}$ denotes the F random variable with the numerator degrees of freedom a and the denominator degrees of freedom b. It follows from Halperin (1961) that, conditionally given Q,

$$\hat{\delta} \sim N\left(\delta, \sigma_{yy.X}(1+Q)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$$

Furthermore,

$$\hat{\sigma}_{yy.X} = \frac{w_{yy} - w_{yX} W_{XX}^{-1} w_{Xy}}{n_1 + n_2 - p - 1} \sim \frac{\sigma_{yy.X}}{n_1 + n_2 - p - 1} \chi^2_{n_1+n_2-p-1}$$

independently of Q. Therefore, conditionally given Q, the pivotal quantity

$$T = \frac{(1/n_1 + 1/n_2)^{-\frac{1}{2}}(\hat{\delta} - \delta)}{\sqrt{\hat{\sigma}_{yy.X}(1 + Q)}} \sim t_{n_1+n_2-p-1}, \qquad (6)$$

where t_m denotes the Student's t variable with degrees of freedom m.

Notice that if samples of data provide evidence against H_0 in Eq. (4), then we conclude that the dissolution profiles are similar. The δ_0 is 10% equivalent to the 50% critical value recommended in the FDA guideline for the f_2 criterion. It is clear from the distribution of $\hat{\delta}$ and Eq. (6) that our problem is essentially testing if a normal mean is contained in a known interval, which has been well addressed in assessing the average bioequivalence of two drugs. A standard approach for the latter problem is due to Schuirmann (1981, 1987). This approach is known as two one-sided tests (TOST), and using this approach we reject the above null hypothesis when

$$\frac{\hat{\delta} + \delta_0}{SE} > t_{n_1+n_2-p-1,1-\alpha} \quad \text{and} \quad \frac{\hat{\delta} - \delta_0}{SE} < -t_{n_1+n_2-p-1,1-\alpha}$$

where

$$SE = \sqrt{\hat{\sigma}_{yy.X}(1 + Q)(1/n_1 + 1/n_2)}$$

In other words, the null hypothesis in Eq. (4) will be rejected if the $(1 - 2\alpha)$ confidence interval $\hat{\delta} \pm t_{n_1+n_2-p-1,1-\alpha}(SE)$ is contained in the interval $(-\delta_0, \delta_0)$. Even though the confidence level is $(1 - 2\alpha)$, the Type I error rates of the above test are always less than or equal to the nominal level α. For more about the Shuirmann's approach, its properties, and other tests for average bioequivalence problem, we refer to Berger and Hsu (1996) and Chow and Liu (2000). If one wants to use the p-value approach, then the null hypothesis in Eq. (4) will be rejected if

$$P_1 = P\left(t_{n_1+n_2-p-1} > \frac{\hat{\delta} + \delta_0}{SE}\right) < \alpha \quad \text{and} \quad P_2 = P\left(t_{n_1+n_2-p-1} < \frac{\hat{\delta} - \delta_0}{SE}\right) < \alpha$$

or

$$p\text{-value} = \max\{P_1, P_2\} < \alpha$$

Remarks

1. Halperin also proposed an unconditional test for the problem considered in his paper. This unconditional test for our present problem is based on the pivotal quantity $[(1/n_1 + 1/n_2)^{-1/2}(\hat{\delta} - \delta)]/\sqrt{\hat{\sigma}_{yy.X}}$, which is distributed as $t_q = t_{n_1+n_2-p-1}(1 + Q)^{\frac{1}{2}}$. Krishnamoorthy and Lu (2005) numerical method can be

readily used to compute the exact percentile points of t_q for a given n_1, n_2, p and α. However, their power comparison studies for the one-sample case showed that the conditional test is slightly more powerful than the unconditional test. Because the conditional test is not only simple to use but also it is more powerful than the unconditional test, we recommend only the conditional test for practical applications.

2. The procedure suggested by Saranadasa (2001) is based on the $(1 - \alpha) \times 100\%$ confidence region for the mean difference vector of the two profiles and assume that the population difference mean vector is of the form $e\delta$, where e denotes the $p \times 1$ vector of ones. Based on the same assumptions considered in this paper, one looks at the set of δ such that:

$$(\bar{u}_d - \delta e)' S^{-1} (\bar{u}_d - \delta e) \leq C, \quad \text{where}$$

$$C = \tfrac{n(n_1+n_2)}{n_1 n_2} \tfrac{p}{n-p+1} F_{p,n-p+1}(\alpha), \text{ and } n = n_1 + n_2 - 2$$

A solution for finding the maximum δ satisfies the above inequality was presented in Saranadasa (2001), and it was used to establish the similarity of

Table 1 Summary statistics of similar dissolution profiles

Figure	Time (min)	Profile 1 (reference)				Profile 2 (test)			
		No. of tablets	Average dissolved (%)	STD (%)	RSD (%)	No. of tablets	Average dissolved (%)	STD (%)	RSD (%)
1a	2	6	22.9	6.6	28.7	6	44.3	9.5	21.3
	5	6	65.9	12.0	18.3	6	75.8	2.9	3.8
	8	6	89.0	2.1	2.3	6	89.0	1.2	1.3
	11	6	93.6	2.5	2.7	6	90.8	1.3	1.4
	15	6	96.9	1.5	1.5	6	94.6	1.5	1.6

$f_2 = 48.5\%$, $|\hat{\delta}| = 1.2$, $T = -5.3$, p-value $= .0009$, maximum IUT upper bound $= 26.0$

	2	6	18.1	1.4	7.8	6	21.8	1.8	8.3
1b	5	6	65.4	7.0	10.7	6	77.8	1.9	2.4
	8	6	85.4	2.0	2.3	6	89.2	1.5	1.7
	10	6	87.8	2.1	2.4	6	92.7	1.5	1.6
	12	6	93.7	5.6	6.0	6	94.8	1.8	1.9
	20	6	97.6	4.7	4.8	6	98.8	2.3	2.3
	60	6	106.3	3.2	3.1	6	107.0	1.4	1.3

$f_2 = 62.7\%$, $|\hat{\delta}| = 3.6$, $T = -14.8$, p-value $< .0001$, maximum IUT upper bound $= 15.3$

1c	10	12	12.0	4.4	37.0	6	13.9	8.4	60.3
	20	12	28.0	6.4	22.8	6	29.1	4.0	13.9
	30	12	45.9	8.8	19.2	6	48.2	5.8	12.0
	45	12	61.5	10.8	17.6	6	74.0	7.7	10.5
	60	12	79.6	9.0	11.3	6	90.1	1.1	1.2

$f_2 = 56.3\%$, $|\hat{\delta}| = 0.24$, $T = -2.4$, p-value $= .0183$, maximum IUT upper bound $= 16.9$

STD: standard deviation; RSD: relative standard deviation $= \left(\tfrac{STD}{Average} \right) \times 100$; T in Eq. 6; maximum IUT upper bound: Max of $|d_k| + c \left(\tfrac{s_k^2}{n} \right)^{1/2}$, $k = 1, 2, \ldots, p$.

two dissolution profiles. In contrast to the above approach, in this article we developed a statistical test for testing hypotheses in Eq. (4) for given specification of δ.

5. PRACTICAL APPLICATIONS

The performance of the proposed procedure was compared with the f_2 criterion on real data examples. Two groups of three-pair dissolution profiles were used. The first group consists of similar dissolution profiles (two lots of the same formulation of a marketed drug), and the second group includes different dissolution profiles of two dosage strengths. The results are summarized in Tables 1 and 2:

In practice, the threshold value for f_2 criterion is allowed as 50% to declare similar dissolution profiles according to the FDA guideline (1). The corresponding mean difference at each time point would be not more than 10%. Therefore, we use 10% mean difference for testing the proposed multivariate test with 5% α-level, 10% mean difference for IUT test with 5% α-level, and 50% threshold for f_2 criterion.

Table 2 Summary statistics of dissimilar dissolution profiles

		Profile 1 (reference)				Profile 2 (test)			
Figure	Time (min)	No. of Tablets	Average dissolved (%)	STD (%)	RSD (%)	No. of tablets	Average dissolved (%)	STD (%)	RSD (%)
2a	2	6	17.2	2.0	11.5	6	44.3	9.5	21.3
	5	6	47.5	2.2	4.7	6	75.8	2.9	3.8
	8	6	75.4	2.0	2.7	6	89.0	1.2	1.3
	11	6	85.3	0.8	0.9	6	90.8	1.3	1.4
	15	6	91.1	1.6	1.7	6	94.6	1.5	1.6

$f_2 = 36.3\%, |\hat{\delta}| = 10.0, T = -.005$, p-value $= .50$, maximum IUT upper bound $= 31.0$

2b	5	6	32.7	6.7	20.4	6	36.0	3.7	10.2
	10	6	65.1	10.6	16.3	6	64.0	3.3	5.2
	20	6	92.4	5.4	5.8	6	89.6	3.9	4.4
	30	6	98.5	4.0	4.0	6	85.8	4.6	5.4
	45	6	101.1	2.7	2.7	6	87.9	3.4	3.9
	60	6	101.9	3.3	3.3	6	85.7	2.3	2.7

$f_2 = 49.6\%, |\hat{\delta}| = 13.6, T = .81$, p-value $= .77$, maximum IUT upper bound $= 17.8$

2c	2	6	17.2	2.0	11.5	6	22.9	6.6	28.7
	5	6	47.5	2.2	4.7	6	65.9	12.0	18.3
	8	6	75.4	2.0	2.7	6	89.0	2.1	2.3
	11	6	85.3	0.8	0.9	6	93.6	2.5	2.7
	15	6	91.1	1.6	1.7	6	96.9	1.5	1.5

$f_2 = 46.9\%, |\hat{\delta}| = 11.1, T = .38$, p-value $= .64$, maximum IUT upper bound $= 23.4$

STD: standard deviation; RSD: relative standard deviation $= \left(\frac{STD}{Average}\right) \times 100$; T in Eq. 6; maximum IUT upper bound: Max of $|d_k| + c\left(\frac{s_k^2}{n}\right)^{1/2}$, $k = 1, 2, \ldots, p$.
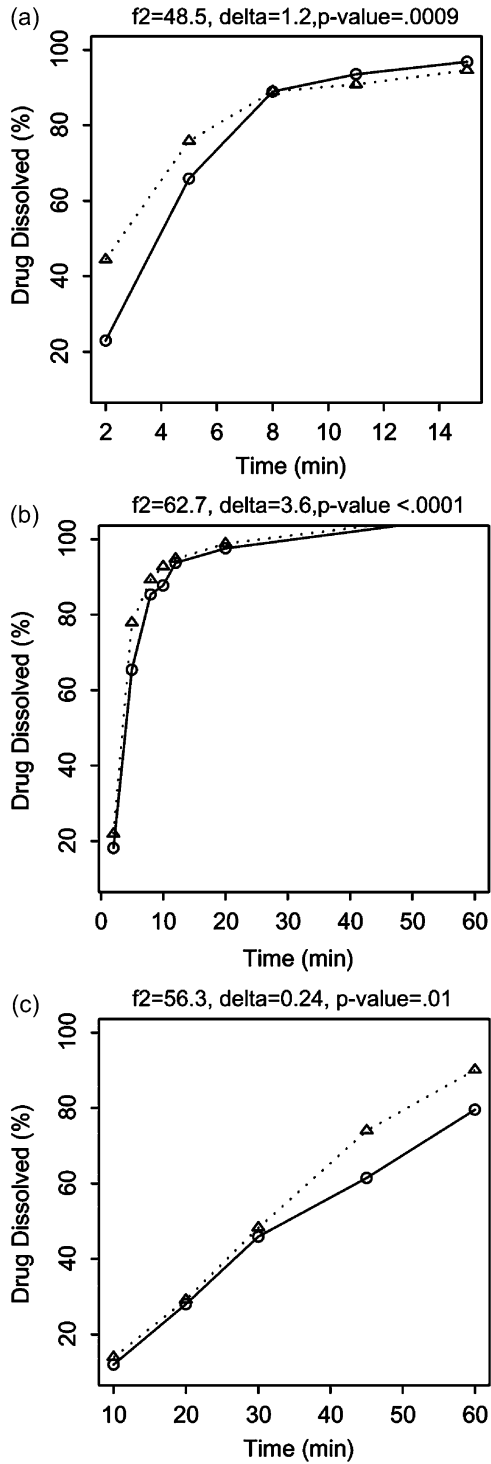
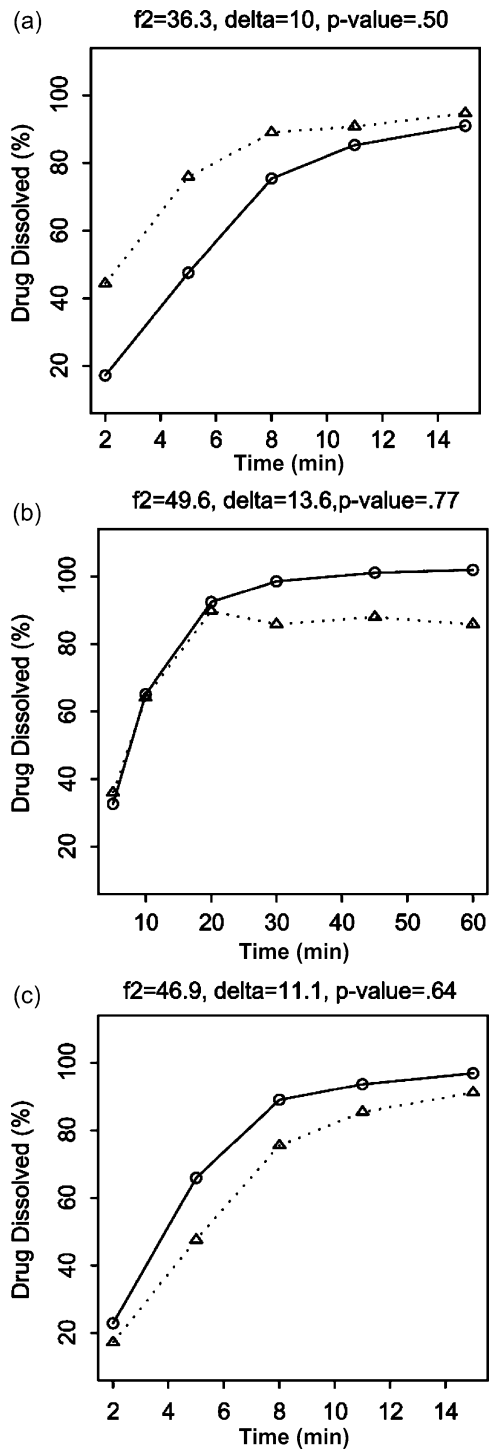**Figure 1** Three similar dissolution profiles.

Figure 2 Three dissimilar dissolution profiles.

Table 1 summarizes the results of three tests for the application of the three similar dissolution profiles. The proposed test correctly classified all three pair profiles as similar (Fig. 1), whereas f_2 criterion misclassified one example (Fig. 1a) as dissimilar profiles. The IUT failed to classify none of the examples as similar profiles even though they were similar.

Table 2 summarizes the classification results of three dissimilar profiles (Fig. 2). All three criteria correctly discriminated all profiles as different dissolution profiles.

6. A SIMULATION STUDY

We carried out a simulation experiment to examine the performance of f_2, Intersection Union test (IUT) and the proposed multivariate test. The multivariate random samples of size 12 were generated from a multivariate normal distribution. The mean vectors for the two profiles were chosen at 5, 10, 15, and 20 minutes following an S-Shape logistic dissolution model of the form:

$$y_t = Q \frac{e^{-b(t-t_{50})}}{1 + e^{-b(t-t_{50})}}$$

with the variance-covariance structure (Σ) of the form:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \rho_{14}\sigma_1\sigma_4 \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 & \rho_{24}\sigma_2\sigma_4 \\ \rho_{31}\sigma_3\sigma_1 & \rho_{32}\sigma_3\sigma_2 & \sigma_3^2 & \rho_{34}\sigma_3\sigma_4 \\ \rho_{41}\sigma_4\sigma_1 & \rho_{42}\sigma_4\sigma_2 & \rho_{43}\sigma_4\sigma_3 & \sigma_4^2 \end{pmatrix}$$

where σ_i is the standard deviation of dissolution values at ith time point ($i = 5$, 10, 15 and 20 minutes) and ρ_{ij} is the correlation coefficient of ijth time point. We chose the variance-covariance parameters, which would mimic the nature of

Table 3 Percent of experiments accepting the equivalency of two dissolution profiles for case 1 (early time point %RSD > 20)

δ	$\rho_{ij} = 0$			$\rho_{ij} = .3$			$\rho_{ij} = .5$			$\rho_{ij} = .7$		
	f_2	SK*	IUT**	f_2	SK*	IUT**	f_2	SK*	IUT**	f_2	SK*	IUT**
0	100	100	97	100	100	96.1	100	100	95.9	100	100	97
2	100	100	93.3	100	100	93.3	100	100	93.8	100	100	94.5
4	100	100	82.8	100	100	83.8	100	100	83.6	100	100	83.6
6	100	100	61.8	99.9	100	61.4	99.7	100	61.3	99.6	100	63.2
8	97.1	97	25.2	93.6	88.1	31.0	92.2	87.9	33.5	90.8	91.3	38.4
9.5	78.3	22.9	1.3	73.8	16.4	6.8	70.1	15.2	10.6	66.8	13.8	14.0
10	41.5	6.1	0.1	48.8	4.7	1	46.0	4.1	2.6	46.8	4.3	4.9
10.5	23.1	1	0	27.7	1.0	0.1	30.3	1.0	0.7	32.9	1.0	2.4
12	0.1	0	0	1.7	0	0	4.3	0	0	6.1	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0

*Saranadasa and Krishnamoorthy.
**Intersection union test.

Table 4 Percent of experiments accepting the equivalency of two dissolution profiles for case 2 (early time point %RSD < 10)

δ	$\rho_{ij} = 0$			$\rho_{ij} = .3$			$\rho_{ij} = .5$			$\rho_{ij} = .7$		
	f_2	SK*	IUT**	f_2	SK*	IUT**	f_2	SK*	IUT**	f_2	SK*	IUT**
0	100	100	100	100	100	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100	100	100	100	100	100
6	100	100	98.1	100	100	98.3	100	100	99	100	100	97.9
8	100	97.9	54.9	99.5	87.1	58.7	99.4	82.7	61.2	99	79.4	66.8
9.5	68.3	24.1	0.9	63.2	20.0	5.1	64.7	17.6	9.0	63.0	20.0	11.9
10	43.5	6.6	0.2	44.4	4.7	1.2	46.3	4.8	2.7	44.3	5.4	4.9
10.5	14.7	0.4	0	23.9	1.0	0.1	22.6	1.7	1.4	26.1	1.5	2.3
12	0	0	0	0.3	0	0	0.8	0	0	1.0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0

*Saranadasa and Krishnamoorthy.
** Intersection union test.

the variability of the typical dissolution experiments. The model parameters and variance-covariance parameters are given for two cases below:

Case 1: $Q = 90.4$, $b = -0.36$, $t_{50} = 6.8$ (time taken to dissolve 50%) $\sigma_1 = 8$, $\sigma_2 = 3$, $\sigma_3 = 2$, $\sigma_4 = 2$ and RSD's are 25.8%, 4.4%, 2.3%, and 2.2% at each time point
$\rho_{ij} = 0$, 0.3, 0.5 and 0.7 for all i, j.

Case 2: $Q = 90.4$, $b = -0.36$, $t_{50} = 6.8$ (time taken to dissolve 50%) $\sigma_1 = 3$, $\sigma_2 = 3$, $\sigma_3 = 2$, $\sigma_4 = 2$ and RSD's 9.7%, 4.4%, 2.3%, and 2.2% at each time point
$\rho_{ij} = 0$, 0.3, 0.5 and 0.7 for all i, j.

One thousand experiments for each configuration were simulated from the respective multivariate normal distribution. The numbers of experiments failing the null hypothesis (i.e., H_a: similar dissolution) by each of three criteria were counted. The $\delta_0 = 10\%$ was used as the acceptable mean difference of two dissolution profiles. The experiments were repeated for mean shift of 0, 2, 4, 6, 8, 9.5, 10, 10.5, 12, and 15% between the two profiles. The results are summarized in Tables 3 and 4.

7. CONCLUDING COMMENTS

In this article, we proposed a multivariate procedure for the case where it is reasonable to make the assumption of multivariate normality of the data. The proposed method was compared with the Intersection Union test as well as f_2 criterion recommended by the FDA guideline. A simulation study was conducted to compare the three procedures for the same data. The results showed that f_2 test and the proposed test have almost same power when $\delta \leq 6\%$, but f_2 test is at the cost of a highly inflated Type I error at the selected mean difference of dissolution profiles that are considered to be similar. This error rate is in general more than 45% on average at 10% mean difference. That is, f_2 test tends to accept different dissolution profiles as similar dissolution profiles with more than 45% of the time at the selected mean difference of dissolution profiles ($\delta = 10\%$).

However, the proposed test maintains the Type I error rate at the selected α-level at the prespecified profile difference, which is considered to be acceptable (10%).

The Intersection Union test showed lower power compared with both f_2 and the proposed test when the data variability was considerably high (see Case 1). In general, the Type I error rates for IUT were under α-level for the most cases.

Because of the inclusion of several time points after the effective completion of dissolution (>85%), the f_2 criterion, as well as the proposed criterion, tends to produce significant results in favor of similarity. Therefore, we also recommend limiting one sampling time point after 85% dissolved for using the proposed criterion [see FDA 1997 guidance for industry]. As is seen in Fig. 1(a), the f_2 criterion and the proposed criterion lead two different conclusions for this case study. The data showed that at 2 minutes tablet-to-tablet variability was very high (RSD > 20%), the observed mean difference was as high as 22% compared with the rest of the time points, and low power (68%) of estimating mean difference within 10% with observed high variability with six tablets. Because the f_2 criterion does not consider tablet-to-tablet variability at each time point (not formulation related) into the calculation of f_2 score, it concludes dissimilar dissolution profile solely on 2 minutes time point difference. On the other hand, the proposed test statistic of the procedure is based on an estimator, which is a weighted average of the mean profile differences at all time points, and the weights are based on the inverse of the estimated variance at each time point. It is reasonable to weight the mean difference by their variance because the variability at each time point is not mostly formulation related but it is due to inherent dissolution method variability.

REFERENCES

Berger, R. L., Hsu, J. (1996). Bioequivalence trials, intersection-union tests and equivalence confidence sets (with discussion). *Statistical Science* 11:283–319.

Chow, S. C., Ki, F. Y. C. (1997). Statistical comparison between dissolution profiles of drug products. *Journal of Biopharmaceutical Statistics* 7(30):241–258.

Chow, S. C., Liu, J. P. (2000). *Design and Analysis of Bioavailability and Bioequivalence Studies* New York: Marcel Dekker.

Eaton, M. L., Muirhead, R. J., Steeno, G. S. (2003). Biopharmaceutical Report, Winter. *Biopharmaceutical Section, American Statistical Association* 11(2):2–7.

Halperin, M. (1961). Almost linearly-optimum combination of unbiased estimates. *Journal of the American Statistical Association* 56:36–43.

Krishnamoorthy, K., Lu, Y. (2005). On combining correlated estimators of the common mean of a multivariate normal distribution. *Journal of Statistical Computation and Simulation*, to appear.

Ma, M. C., Lin, R. P, Liu, J. P. (1999). Statistical evaluation of dissolution similarity. *Statistica Sinica*, 9:1011–1027.

Ma, M. C., Wang B. B. C., Liu, J. P., Tsong, Y. (2000). Assessment of similarity between dissolution profiles. *Journal of Biopharmaceutical Statistics* 10:229–249.

Moore, J. W., Flanner, H. H. (1996). Mathematical comparison of dissolution profiles. *Pharmaceutical Technology* 20:64–74.

Saranadasa, H. (2001). Defining similarity of dissolution profiles through hotelling T^2 statistic. *Pharmaceutical Technology* 24:46–54.

Sathe, P. M., Tsong, Y., Shah, V. P. (1996). In vitro dissolution profile comparision: statistics and analysis, model dependent approach. *Pharmaceutical Research* 13(12):1799–1803.

Schuirmann, D. J. (1981). On hypothesis testing to determine if the mean of a normal distribution is contained in a known interval. *Biometrics* 37:617.

Schuirmann, D. J. (1987). Comparison of two one-sided procedures and power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics* 15:657–680.

Shah, V. P., Tsong, Y., Sathe, P., Liu, J. P. (1998). In vitro dissolution profiles comparision—statistics and analysis of the similarity factor f_2. *Pharmaceutical Research* 15(6): 889–896.

Tsong, Y., Hammerstrom, T., Chen, J. J. (1997). Multipoint dissolution specification and acceptance sampling rule based on profile modeling and principal component analysis. *Journal of Biopharmaceutical Statistics* 7(3):423–439.

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

US Food and Drug Administration Guidance. (1995). Immediate Release Solid Oral Dosage Form: Scale-Up and Postapproval Changes, Chemistry, Manufacturing and Controls, In Vitro Dissolution Testing, and In Vivo Bioequivalence Documentation (SUPAC-IR), Rockville, MD.

US Food and Drug Administration, Guidance for Industry. (1997). Dissolution Testing of Immediate release Solid Oral Dosage Forms. Rockville, MD.

US Food and Drug Administration Guidance. (1997). Modified Release Solid Oral Dosage Form: Scale-Up and Postapproval changes, Chemistry, Manufacturing and Controls, In Vitro Dissolution Testing, and In Vivo Bioequivalence Documentation (SUPAC-MR), Rockville, MD.

Wang, W., Hwang, J. T., DasGupta, A. (1999). Statistical tests for multivariate bioequivalence. *Biometrika* 86:394–402.