MedCrave
Step into the World of Research

Biometrics & Biostatistics International Journal

Research Article

Open Access

CrossMark

Bootstrap confidence intervals for dissolution similarity factor f_2

Abstract

Parametric and non-parametric bootstrap methods are used to investigate the statistical properties of the dissolution similarity factor f_2. The main objective of this study is to compare the results obtained by these two methods. We estimate characteristics of the sampling distribution of $\hat{f}_2$ statistic under these methods with various bootstrap sample sizes using Monte Carlo simulation. A number of bootstrap confidence interval (CI) construction techniques are used to determine a 90 % CI for the true value of f_2 under both parametric and non-parametric schemes. The bootstrap sampling distributions of $\hat{f}_2$ under both schemes are found to be approximately symmetrical with a non-zero excess of kurtosis. Non-parametric bootstrap confidence intervals for f_2 perform better than those obtained from parametric methods. The Bias corrected (BC) and accelerated bootstrap percentile (BC_a) confidence interval method produce more precise two-sided confidence intervals for f_2 compared to other methods

Keywords: dissolution profiles, bootstrapping, confidence interval, bias-corrected and accelerated bootstrap percentile confidence interval

Volume 7 Issue 5 - 2018

Mohammad M Islam,[1] Munni Begum[2]

[1]Department of Mathematics, Utah Valley University, USA

[2]Department of Mathematical Sciences, Ball State University, USA

Correspondence: Mohammad M Islam, Department of Mathematics, Utah Valley University, 800 W University Pkwy, Orem, UT 84058, USA, Tel 1801 8636 430, Email mislam@uvu.edu

Received: August 16, 2018 | **Published:** September 18, 2018

Introduction

In pharmaceutical studies for solid and oral drugs, it is important to compare a test drug to a reference drug using average dissolution rates over time. The purpose of dissolution testing is to develop a new formulation, to ensure quality control, and to assess stability and reproducibility of the immediately released solid oral drug.[1–3] Assessment of dissolution profiles for two drugs, in vitro, provides the waiver for in-vivo assessment.

The United States Food and Drug Administration (FDA) requires similarity tests for the dissolution profiles of two drugs under consideration when there are post drug-approval changes. Such changes include change of manufacturing sites, change in formulations, and change in component and composition. Despite the post-approval changes, two drugs are similar with respect to their dissolution rates if the test (post-approval) has the same (equivalent) dissolution performance as the reference (pre-change).

In order to assess drug dissolution profiles, both model-dependent and model-independent methods are used. In a model-dependent approach, an appropriate mathematical model is selected to describe the dissolution profiles of the two drugs. The model is then fit to the data and confidence intervals for the model parameters are constructed. These confidence intervals are then compared with the specified similarity region. Commonly used model-dependent methods to fit the dissolution profiles include Gompertz,[4] Logistic,[5] Weibull,[6] probit and sigmoid models.[7,8] Model-dependent methods have some limitations. For example, selecting an appropriate model, and interpreting its parameters are difficult when the dissolution profiles for the two drugs follow different models.

To overcome the limitations of model-dependent approaches, model-independent approaches such as difference factor f_2, similarity factor f_2, analysis of variance, split plot analysis, repeated measure analysis, Hotelling T^2, principal component analysis[1] and first order autoregressive time series analysis are used. Among these methods, analysis of variance, split plot analysis assume that dissolution data are independent over time. These two methods are not appropriate in many cases as data are not independent. As an alternative, Tsong et al.,[9] proposed Hotelling f_2 statistic to construct a 90% confidence region for the difference in dissolution means of two batches of the reference product at two time points. This confidence region is then compared with a pre-specified similarity region.

Of all the model-independent approaches, the US FDA recommends only f_2 [10] to study similarity between two drug dissolution profiles under consideration. Although this similarity factor f_2 is used to assess global similarity of dissolution profiles, and it does not require any assumption regarding data generating process, using point estimate of f_2 in comparing two drug dissolution profiles is not appropriate if there is substantial variation from batch to batch. In this case, it is necessary to construct the confidence intervals for f_2. Construction of confidence interval for $\hat{f}_2$ depends on the standard error of its estimator $\hat{f}_2$. Since there is no closed form formula for the standard error of $\hat{f}_2$, and hard to derive analytically, we approximate the standard error of $\hat{f}_2$ by deriving sampling distribution of $\hat{f}_2$ using parametric and non-parametric bootstrap methods. Then the approximated standard error of $\hat{f}_2$ is used to construct bootstrap confidence intervals for f_2.

The organization of this paper is as follows: in section 2 we present basic characteristics of drug dissolution data used in our study and chi-square plot for assessing the normality of the underlying population of the data. Section 3 discusses the statistical framework used in dissolution testing and gives an outline how two drugs are considered to be similar in terms of dissolved drug ingredients into the media. In section 4, bootstrap methods are briefly discussed and related confidence intervals for the true value of $\hat{f}_2$ are presented. Section 5 discusses the results of our study and section 6 concludes the paper.

Dissolution data

We consider the standard dissolution data discussed by Chow & Liu,[11] Tsong[12,13] to assess the various characteristics of f_2. The

Summary measures of the reference and the test drug dissolution data are presented in Table 1.

Table1 Summary measures of test and reference drug dissolution data

Time (Hour)							
	1	2	3	4	6	8	10
Test Drug	-	-	-	-	-	-	-
Mean	36.5	50.08	62.17	67.92	79.33	86.42	92
St. Deviation	1.38	2.27	1.47	3	2.53	3.73	2.41
Range	5	9	5	12	8	12	8
Reference Drug	-	-	-	-	-	-	-
Mean	45.08	54	62.5	67.08	74.75	80.25	85.33
St. Deviation	3.2	3.41	3.32	3.75	3.93	4.2	4.5
Range	12	11	14	15	15	16	17
Mean Difference	-8.58	-3.92	-0.33	0.83	4.58	6.17	6.67

The observed difference between mean dissolution rate factors, f_2 for the test and the reference drug at different time points are less than 10 percent. The standard deviation of the dissolution rate factor at different time points for the test and the reference drug are also less than 10 percent. The mean differences between the two drugs are wider at the starting time points than in the mid-time points. Figure 1 shows dissolution profiles for the test and the reference drugs.

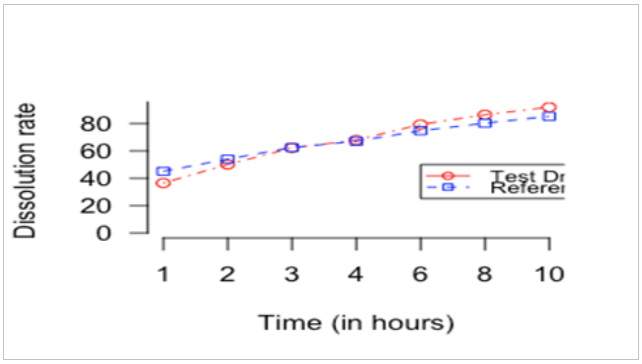

Figure 1 Dissolution profiles for the test and the reference drugs

In order to perform parametric bootstrapping using the above data, we need to know the parametric form for the distribution of the population from which these sample dissolution factors are drawn. In particular we check if the sample dissolution factors are drawn from multivariate normal distribution. To check normality, we examine the underlying distribution of the data using a chi-square plot and a normality goodness of fit test. Because the observations from the same tablets across time are related and the observations across the tablets at a fixed time point are independent, the dissolution data used in this article are considered to be a realization of multivariate observations. To check whether the dissolution data we consider for our study come from the multivariate normal distribution, we calculate statistical distance measures and use them to construct a chi-square plot under the normality assumption (Figure 2).

Since the points in plots are not on a straight line, we say that the data do not follow multivariate normal distributions. We also use the formal correlation test to measure the straightness of the Q-Q plot. The values of the correlation coefficient for the Q-Q plot for the test and the reference drug dissolution data are 0.95 and 0.91 respectively. At 5% level of significance the tabulated value of the correlation coefficient for sample size of $n = 12$ is 0.9298. For the test drug dissolution data, the normality assumption is reasonable but for the reference drug, normality assumption is off slightly.

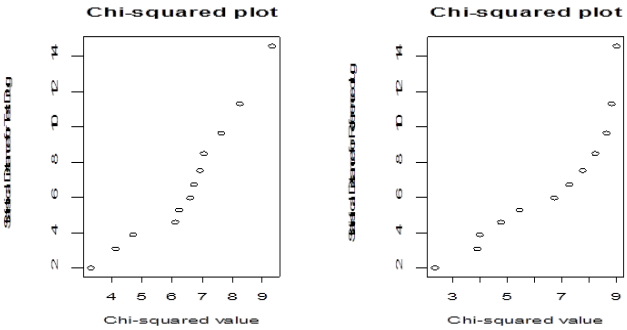

Figure 2 Left panel: Chi-square plot for the test drug. Right panel: chi-square plot for the reference drug

Statistical methods for drug dissolution

Let y_{ijt} be the percentage of drug dissolved in a media at time point t from the tablet i for drug j. Then the statistical model for the drug dissolution percentage can be written as $y_{ijk} = \mu_{jt} + \varepsilon_{ijt}$, $t = 1,2,...,T$; $i = 1,2,...,n$; $j = 1,2$. Here μ_{jt} is the population mean over tablets at time t for drug j and ε_{ijt} has mean 0. Since the dissolution percentage is measured over time from the same tablet of the drug, the measurements are dependent. However, it is reasonable to assume that the n vectors $\mathbf{y}_{1j}, \mathbf{y}_{2j}, ..., \mathbf{y}_{nj}, j = 1,2$ are independent, as these are replications across tablets in the population. The dissolution profiles (Figure 1) of the test drug and the reference drug are considered to be similar if and only if the population means vector for the test drug is in some neighborhood of the population mean vector for the reference drug. A rectangular similarity measure recommended and required by FDA is used to assess if two drugs are similar. This similarity measure is, $|\mu_{1t} - \mu_{2t}| \le \Delta$ where Δ a specified number is. Generally FDA recommends that the specified number is 10 for all time point. A similarity measure f_2, based on the rectangular measure and recommended by FDA, is discussed in Section 3.1.

Similarity factor f_2

Moore et al.,[10] developed a similarity factor, f_2, for testing dissolution profiles of a test and a reference drug.

The f_2 similarity factor is defined as

$$f_2 = 50 \log\left\{100\left(1 + D^2\right)^{-\frac{1}{2}}\right\},$$

where $D^2 = \frac{1}{T}\sum_{t=1}^{T}\left(\mu_{1t} - \mu_{2t}\right)^2$ and μ_{jt} is population mean dissolution rates over time $t\left(t = 1,2,...,T\right)$ and for the jth drug $\left(j = 1,2\right)$. D^2 is a squared distance from the population mean vector of the test drug to the population mean vector of the reference drug. Since dissolution measurements are expressed as percent, D^2 ranges from 0 to 100^2.

Citation: Islam MM, Begum M. Bootstrap confidence intervals for dissolution similarity factor f_2. *Biom Biostat Int J.* 2018;7(5):397–403.
DOI: 10.15406/bbij.2018.07.00237

The similarity factor f_2 is a monotone decreasing function of D^2 with a maximum of 100 when $D^2 = 0$ (two dissolution profiles are the same), and a minimum of 0, when $D^2 = 100^2$.

A value of f_2 in the range of 50 to 100 ensures the similarity or equivalence of two dissolution profiles. When the rectangular similarity measure (adopted by FDA) is $|\mu_{2t} - \mu_{1t}| \le 10$ for all time points, then f_2 is very close to 50. So the similarity region in the range of 50 to 100 indicates the similarity of two drugs.

This similarity factor works well when the following conditions are met: (i) there is a minimum of three time points, (ii) there are 12 individual values for each time point for each formulation, (iii) no more than one mean value is greater than 85% dissolved for each formulation, and (iv) the standard deviation of the mean of any product is less than 10% from the second to last time points.

Bootstrap methods

Bootstrapping[14,15] is a computer-intensive approach to statistical inference. It is based on the sampling distribution of a statistic obtained by resampling from the data with replacement. When it is hard to derive the exact sampling distribution of certain statistics and their characteristics, bootstrap methods are used to approximate them. The characteristics include standard error, bias, skewness, critical values, mean squared error, and others. To derive an exact sampling distribution of a statistic of interest, the underlying population distribution from which sample is drawn has to be known. Sometimes even though the underlying distribution is known, derivation of the exact sampling distribution for certain statistic is not possible or is very complex. In such case, bootstrap methods allow estimating or approximating the sampling distributions of these statistics. The bootstrap approach does not require knowledge of the data generating process but uses the sample information only. The idea behind bootstrapping is that the use of sample information as a "proxy population". One takes samples with replacement from the original sample and calculates the statistic of interest repeatedly. This leads to a bootstrap sampling distribution. This sampling distribution is used to measure the estimator's accuracy and helps to set approximate confidence intervals for certain population parameters.

We use two types of bootstrap methods, parametric and non-parametric to determine the sampling distribution of the statistic $\hat{f}_2$ and its characteristics. Using both techniques we construct 90% confidence intervals for f_2. We briefly describe both methods as follows.

Let $X_1, X_2, ..., X_n$ be independent and identically distributed random variables from an unknown distribution F. F is estimated using the empirical distribution $\hat{F}_n$. Repeated samples are taken from the estimated empirical distribution $\hat{F}_n$. Then the statistic of interest is calculated using each bootstrap samples, giving a set of bootstrap values for the desired statistic. Using the bootstrap values of the statistic, the estimated distribution function and its properties are calculated. This approach is called non-parametric bootstrapping.

The parametric bootstrap, on the other hand, assumes that F is known except for its parameters. F is approximated by estimating the parameters with the sample observations. Then from the approximated distribution $\hat{F}_n$, repeated samples are taken. The values of the statistic of interest are calculated using these bootstrap samples. These bootstrap values of the statistic are used to derive the desired measures. Under both schemes, the distribution of $\hat{f}_2$ can be estimated by using the bootstrap with the Monte Carlo approximation as follows.

$$F(x) = \frac{1}{B}\sum_{b=1}^{B} I\left\{f_2^* \le x\right\},$$

where f_2^* is the value of the f_2 based on the bth bootstrap sample and $F(x)$ is a bootstrap estimator of the distribution function of $\hat{f}_2$ based on the data $X_1, X_2, ..., X_n$. The bootstrap histogram for $\left\{f_2^*, b = 1, 2, ..., B\right\}$ can be used to estimate the density of $\hat{f}_2$. The expected value, variance, skewness, kurtosis, and bias of the bootstrap sampling distribution of $\hat{f}_2$ are estimated from $F(x)$. In order compute them, we first take B independent samples $\left\{X_1^{*,b}, X_2^{*,b}, X_3^{*,b}, ..., X_n^{*,b}\right\}$, $b = 1, 2, ..., B$ and approximate them by

$$\bar{\hat{f}}_2 = \frac{1}{B}\sum_{b=1}^{B} f_{2,b}^*, \quad \upsilon_{\text{boot}}^{(B)} = \frac{1}{B}\sum_{b=1}^{B}\left(f_{2,b}^* - \bar{\hat{f}}_2\right)^2,$$

$$sk_{\text{boot}}^{(B)} = \frac{\frac{1}{B}\sum_{b=1}^{B}\left(f_{2,b}^* - \bar{\hat{f}}_2\right)^3}{\left[\upsilon_{\text{boot}}^{(B)}\right]^{\frac{3}{2}}}, \quad k_{\text{boot}}^{(B)} = \frac{\frac{1}{B}\sum_{b=1}^{B}\left(f_{2,b}^* - \bar{\hat{f}}_2\right)^4}{\left[\upsilon_{\text{boot}}^{(B)}\right]^2}$$

and $b_{Boot}^{(B)} = \bar{\hat{f}}_2 - \hat{f}_2$;

here $\bar{\hat{f}}_2$, $\upsilon_{\text{boot}}^{(B)}$, $sk_{\text{boot}}^{(B)}$, $k_{\text{boot}}^{(B)}$, and $b_{Boot}^{(B)}$ are Monte Carlo bootstrap estimator for mean, variance, skewness, kurtosis and bias of the sampling distribution of $\hat{f}_2$ respectively. In section 4, we construct bootstrap confidence intervals for f_2 using a number of available bootstrap confidence interval methods.

Bootstrap confidence intervals

An observed value of $\hat{f}_2$ is used to assess whether two drugs (test and reference) are similar or not with respect to their dissolution profiles. This value is compared with the specifications given by the FDA in order to decide if the two drugs are similar. However, due to sampling variation, it is not reasonable to assess the dissolution similarity of two drugs by directly comparing $\hat{f}_2$ with the specification limits. Rather one can make a decision of dissolution similarity by constructing a 90% confidence interval for the population parameter $f_2 = E(\hat{f}_2)$. In this section we use parametric and non-parametric bootstrap methods discussed in subsection 3.2 to construct the confidence intervals for f_2.

A detailed discussion on different types of bootstrap confidence intervals can be found in Chernick,[16] Davison,[17] DiCiccio,[18] Efron.[19] Here we review different types of bootstrap procedures used to construct confidence intervals for the parameter of interest briefly. For notational convenience, we denote the similarity parameter f_2 as θ and $\hat{\theta}$ as its estimate.

The standard bootstrap confidence interval is given by

$$\left[\hat{\theta} - z^{\left(\frac{\alpha}{2}\right)}\hat{se}_{(\hat{\theta}^*)}, \hat{\theta} + z^{\left(1-\frac{\alpha}{2}\right)}\hat{se}_{(\hat{\theta}^*)}\right]$$

where $\hat{se}_{(\hat{\theta}^*)}$ is the bootstrap standard error of the estimator $\hat{\theta}$,

Citation: Islam MM, Begum M. Bootstrap confidence intervals for dissolution similarity factor f_2. *Biom Biostat Int J.* 2018;7(5):397–403.
DOI: 10.15406/bbij.2018.07.00237

and $z^{\left(\frac{\alpha}{2}\right)}$ is the $100.\frac{\alpha}{2}th$ quantile of the standard normal distribution. In percentile interval method of bootstrapping, B bootstrap estimates $\hat{\theta}_b^*$ $(b=1,2,...,B)$ are generated. Then these bootstrap estimates are arranged in ascending order. If we denote $F\left(\hat{\theta}^*\right)$ as the cumulative distribution function of $\hat{\theta}^*$, then a 90% percentile interval is defined by

$$\left(\hat{\theta}_b^{*\left(\frac{\alpha}{2}\right)}, \hat{\theta}_b^{*\left(1-\frac{\alpha}{2}\right)}\right) = \left[\hat{F}^{-1}\left(\frac{\alpha}{2}\right), \hat{F}^{-1}\left(1-\frac{\alpha}{2}\right)\right]$$

where $\alpha = 5\%$ and $\hat{\theta}^{*\left(\frac{\alpha}{2}\right)}$ indicates the $100.\frac{\alpha}{2}th$ percentile of B bootstrap replications.

Although the computation is straightforward, this method does not work well when the sampling distribution of $\hat{\theta}$ is skewed or $\hat{z}_0$ is biased.[20,21]

In the presence of skewness and bias, the percentile method can be improved by an adjustment to the percentile method. This bias adjusted and corrected percentile interval is known as bias corrected percentile interval method (BC).[22] In the bias-corrected method, the observed amount of difference between the median of the bootstrap estimate $\hat{\theta}^*$ and the observed estimate from the original sample is defined as bias. The bias-correction constant estimate, denoted by $\hat{z}_0$, is defined as

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\hat{\theta}^* < \hat{\theta}}{B}\right),$$

where Φ^{-1} is the inverse function of a standard normal cumulative distribution function. Then, a $100(1-\alpha)$ percent bias-corrected percentile confidence interval for θ is given by $\left[\hat{\theta}^{*(\alpha_1)}, \hat{\theta}^{*(\alpha_2)}\right]$, where

$$\alpha_1 = \Phi\left(2z_0 + z^{\left(\frac{\alpha}{2}\right)}\right)$$

$$\alpha_2 = \Phi\left(2z_0 + z^{\left(1-\frac{\alpha}{2}\right)}\right).$$

Here Φ is the standard normal cumulative distribution function and $z^{\left(\frac{\alpha}{2}\right)}$ is the $100\frac{\alpha}{2}$th percentile point of the standard normal distribution. Although the bootstrap bias correction improves the bootstrap percentile method with taking the bias into account, this method does not work well in some cases.[21]

Efron[22] introduced a further improved bootstrap method that corrects the bias due to the non-normality and also accelerates convergence to a solution. The method corrects the rate of change of the normalized standard error of $\hat{\theta}$ relative to the true parameter θ. It takes into account the skewness in the distribution along with the bias of the estimator. This method is called bias-corrected and accelerated (BC_a) percentile method. Chernick et al.,[16] show that for small sample sizes BC_a may not work as percentile method because the bias and acceleration constant must be estimated and the sample size is not large enough for asymptotic advantage of BC_a to hold.

The BC_a confidence interval for θ is

$$\left[\hat{\theta}^{*\left(\alpha_1^*\right)}, \hat{\theta}^{*\left(\alpha_2^*\right)}\right],$$

where

$$\alpha_1^* = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{\left(\frac{\alpha}{2}\right)}}{1 - a\left(\hat{z}_0 + z^{\left(\frac{\alpha}{2}\right)}\right)}\right) \text{ and}$$

$$\alpha_2^* = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{\left(1-\frac{\alpha}{2}\right)}}{1 - a\left(\hat{z}_0 + z^{\left(1-\frac{\alpha}{2}\right)}\right)}\right),$$

where Φ is the standard normal cumulative distribution function and $z^{(.)}$ is the percentile of standard normal distribution.

The bias correction term $\hat{z}_0$ is calculated by $z_0 = \Phi^{-1}\left(\frac{\#\left(\hat{\theta}^* < \hat{\theta}\right)}{B}\right)$, and the acceleration constant a is :

$$a = \frac{\sum_{i=1}^{n}\left(\hat{\theta}_{(.)} - \hat{\theta}_{(i)}\right)^3}{6\left[\sum_{i=1}^{n}\left(\hat{\theta}_{(.)} - \hat{\theta}_{(i)}\right)^2\right]^{\frac{3}{2}}},$$

where $\hat{\theta}_{(.)}$ and $\hat{\theta}_{(i)}$ are the average and ith jackknife estimate of the parameter.

The problem arising from the skewness in the sampling distribution of $\hat{\theta}$ can also be handled by an alternative method called Bootstrap-t method.[22, 26, 27] The bootstrap-t method is defined by the pivotal quantity $t^* = \frac{\hat{\theta}^* - \hat{\theta}}{se_{\hat{\theta}^*}}$, where $\hat{\theta}^*$ and $se_{\hat{\theta}^*}$ are the bootstrap estimator and its standard error. Since the standard error of $\hat{\theta}$ is not known, it is estimated by Monte Carlo simulation. However, this simulation requires nested bootstrapping. For each bootstrap sample, we calculate t^* and these resulting t^*'s are placed in ascending order and select $100\left(\frac{\alpha}{2}\right)$th and $100\left(1-\frac{\alpha}{2}\right)$th percentile values of t^*. Then, $100(1-\alpha)$ percent bootstrap-t confidence for $\hat{\theta}$ is

$$\left[\hat{\theta} - t^*_{1-\frac{\alpha}{2}} se_{\hat{\theta}^*}, \hat{\theta} - t^*_{\frac{\alpha}{2}} se_{\hat{\theta}^*}\right]$$

where $\hat{\theta}$ is estimate of the parameter θ from the original sample and $se_{\hat{\theta}^*}$ is the bootstrap standard error.

Results and discussion

Properties of the distribution of $\hat{f}_2$

In this section, we examine the properties of the bootstrap sampling distribution of $\hat{f}_2$ by using both non-parametric and parametric bootstrap sampling. To generate the bootstrap samples

Citation: Islam MM, Begum M. Bootstrap confidence intervals for dissolution similarity factor f_2. *Biom Biostat Int J.* 2018;7(5):397–403.
DOI: 10.15406/bbij.2018.07.00237

with non-parametric and parametric bootstrap methods discussed in subsection 3.3. the following algorithms are employed: (a) For Non-parametric bootstrapping: $B(b=1,2,...,B)$ independent sample with replacement from the observed $\mathbf{y}_{1j},\mathbf{y}_{2j},...,\mathbf{y}_{nj}, j=1,2$ are drawn and for each bootstrap sample $\hat{f}$, which is the estimate of f_2 defined in Subsection 3.1, is calucated; (b) For Parametric bootstrappping: B independent samples are drawn from $N_T(\hat{\mu}_j,\hat{\Sigma}_j), j=1,2$, where $\hat{\mu}=(\hat{\mu}_{j1},\hat{\mu}_{j2},...,\hat{\mu}_{jT})^T$ and $\hat{\Sigma}_j$ of $T\times T$ covariance matrix are moment estimates of μ_j and Σ_j and for each bootstrap sample $\hat{f}_2$, the estimate of f_2 is calculated.

The histograms and Q-Q plots are constructed using the bootstrap values of $\hat{f}_2$ generated by both methods and are shown in Figure 3 & Figure 4 respectively.

The Bootstrap parametric and non-parametric sampling distributions of $\hat{f}_2$ shown in the left and right panels of Figure 2 are almost symmetrical. However, the non-parametric sampling distribution of the similarity factor $\hat{f}_2$ is more symmetrical than that of the parametric one. In addition, the bootstrap parametric sampling distribution of $\hat{f}_2$ is wider than the non-parametric bootstrapping. The percentile confidence methods work well if the underlying probability distribution from which samples are drawn is symmetric and the distribution of statistics is also symmetric. The reliability of the confidence interval for true parameter f_2 by bootstrap method relies upon the symmetrical pattern of the sampling distribution of $\hat{f}_2$. The Q-Q plot of the sampling distribution of $\hat{f}_2$ generated by a parametric bootstrap method, given in the left panel of Figure 3 confirms normality better than the Q-Q plot in the right panel generated by non-parametric bootstrap method. However, Q-Q plots do not confirm the normality assumption. So we apply more rigorous statistical test to verify the normality of the data. A commonly used normality test is Jargue-Bera test [3], which is based on skewness and kurtosis. In what follows we present some characteristics of the sampling distribution of $\hat{f}_2$ obtained by both methods and the results of Jargue-Bera test. To assess basic properties of $\hat{f}_2$, we carry out empirical simulation study by Monte Carlo method. We estimate the characteristics of the sampling distribution defined in Section 3.2 with Monte Carlo size $B=3000$. The simulation average (ME) of the statistics, and the coefficient of variation (CV), the ratio of the standard deviation of the statistic over the absolute value of ME based on 100 simulation replications, are presented in Table 2.

Table 2 Summary measures of sampling distribution of $\hat{f}_2$ under parametric and non-parametric bootstrap method

Method	ME	CV
Non-parametric		
Mean	63.2	0.0009
Variance	4.24	0.0427
Coefficient of Skewness	0.01	0.4776
Coefficient of Kurtosis	3.25	0.0619
Bias	-0.43	0.1778
Parametric		
Mean	63.15	0.0012
Variance	4.46	0.0476
Coefficient of Skewness	0.14	0.5938
Coefficient of Kurtosis	3.11	0.0619
Bias	-0.43	0.1776

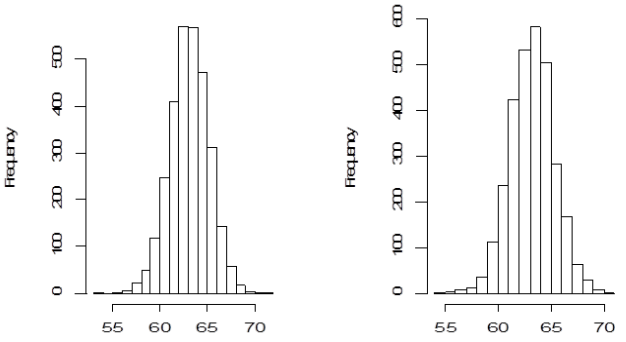

Figure 3 Left panel: Histogram of 3000 parametric bootstrap of $\hat{f}_2$. Right panel: Histogram of 3000 non-parametric bootstrap of $\hat{f}_2$.

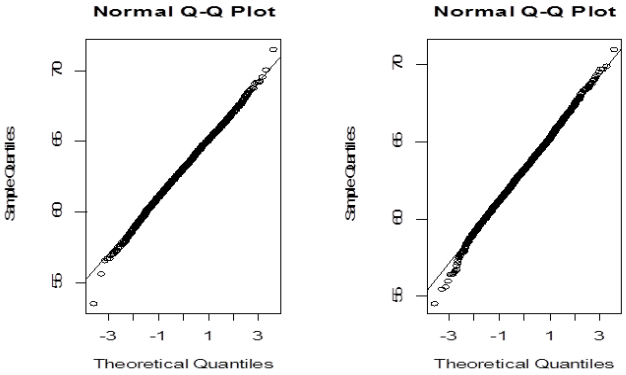

Figure 4 Left panel: Q-Q plot for the values of $\hat{f}_2$ generated by parametric bootstrap method. Right panel: Q-Q plot for the values of $\hat{f}_2$ generated by non-parametric bootstrap method.

Both bootstrap estimators of the expected value of the sampling distribution $\hat{f}_2$ are downward-biased. Non-parametric estimators are better than those obtained in parametric method in terms of variance and skewness. The coefficients of skewness in both procedures indicate that the sampling distribution is almost symmetric but slightly positively skewed. The coefficients of kurtosis under both procedures are slightly more than 3. The sampling distribution of $\hat{f}_2$ is approximately normal. We calculate Jargue-Bera test statistic to check the normality of the distribution of $\hat{f}$. For a symmetric distribution, the third moment about mean (μ_3) and coefficient of skewness (τ) are equal to zero. The normal distribution is characterized with $\tau=0$ and coefficient of kurtosis, $k=3$. A joint test of $\tau=0$ and $k=3$ is often used as a test of normality. Jargue & Bera[26] proposed a statistic to test the normality of a distribution. Their proposed test statistic under the normality assumption is

$$JB=B\left(\frac{\tau^2}{6}+\frac{(k-3)^2}{24}\right)\tilde{\chi}_2^{\,2},$$

where B is the number of the bootstrap samples. We can use this statistic to test the normality of the distribution of $\hat{f}_2$. We have $JB=7.86$ and 11.31 for parametric and non-parametric sampling distribution of $\hat{f}_2$ respectively. These are highly insignificant (

Citation: Islam MM, Begum M. Bootstrap confidence intervals for dissolution similarity factor f_2. *Biom Biostat Int J.* 2018;7(5):397–403.
DOI: 10.15406/bbij.2018.07.00237

$\chi_2^2(0.05) = 5.991$) and ($\chi_2^2(0.01) = 9.21$. Thus we may conclude that the distribution of f_2 is not normal under both procedures, and we apply bootstrap algorithms to construct CIs for f_2 .

Confidence Intervals for f_2

For each bootstrap method of sampling, bootstrap-t, percentile, bias-corrected, and bias-corrected and accelerated confidence intervals for the parameter $\hat{f}_2$ are constructed and presented in Table 3.

Table 3 Non-parametric and Parametric Bootstrap Confidence Intervals for $\hat{f}_2$

Method	500 Bootstraps		1000 Bootstraps		1500 Bootstraps		2000 Bootstraps		2500 Bootstraps	
Non-parametric	**Mean**	**CI**	**Mean**	**CI**	**Mean**	**CI**	**Mean**	**CI**	**Mean**	**CI**
Bootstrap-t	63.17	(60.75-67.59)	63.21	(60.66-67.40)	63.21	(60.67-67.29)	63.19	(60.63-67-04)	63.21	(60.73-67.24)
BP		(59.88,66.48)		(59.83,66.37)		(59.89,66.44)		(59.75,66.55)		(59.74,66.48)
BC		(60.72-67.49)		(60.91-67.49)		(60.66-67.18)		(60.75-67.24)		(60.78-67.30)
BC_a		(60.86,66.99)		(60.58,67.22)		(60.70,67.36)		60.64,67.29)		(60.65,67.31)
Parametric										
Bootstrap-t	63.12	(60.88-67.30)	63.09	(60.76-67.49)	63.13	(60.68-67.64)	63.19	(60.54-67.58)	63.15	(60.73-67.63)
BP		(59.51-66.49)		(59.74-66.80)		(59.54-66.47)		(59.71-66.48)		(59.64-66.49)
BC		(60.52-67.22)		(60.41-67.29)		(60.42-67.01)		(60.45-67.14)		(60.33-67.14)
BC_a		(60.61-67.27)		(60.46-67.32)		(60.44-67.04)		(60.49-67.23)		(60.36-67.19)

The observed value of $\hat{f}_2$ for the original dissolution data is 63.58. At 10% average distance at all time-points the similarity criterion is 50. Since the point estimate of f_2 , $\hat{f}_2 = 63.58$ is greater than the criterion value of 50, two drugs are considered to be same in terms of average dissolution data. The table 3 shows the 90% confidence interval for $E(\hat{f}_2)$ with the bootstrap replications $500 : 500 : 2500$.

Under parametric and non-parametric bootstrap sampling schemes and all the bootstrap CI methods, the 90% lower confidence interval for $E(\hat{f}_2)$ is greater than the similarity criterion value 50. This indicates that two drugs are similar.

However, under both parametric and non-parametric approaches, the percentile confidence interval is wider than the other bootstrap confidence intervals. This method, however, does not incorporate the skewness of the sampling distribution of $\hat{f}_2$. The BC_a method corrects the bias and skewness in the sampling distribution of statistic. In the setting of both parametric and non-parametric procedures BC_a gives the shortest confidence interval for $E(\hat{f}_2)$. The accuracy of the parametric and non-parametric bootstrap approximate confidence intervals for $E(\hat{f}_2)$ cannot be accessed directly just by eyeballing. To see which method works well in our situation, we perform empirical comparisons of these bootstrap confidence intervals in the next section.

Empirical comparisons

In this section we examine and compare Bootstrap-t, BP, BC, and BC_a confidence sets using simulation approach. The Monte Carlo simulation with size 500 is used to calculate the simulation average of a confidence bound (AV), and the simulation estimates of the expected length of a two-sided confidence interval (EL). This simulation study is performed under both non-parametric and parametric bootstrap sampling schemes.

Table 4 shows average left and right endpoints of the confidence intervals constructed by various methods and the expected length of two sided confidence intervals for $E(\hat{f}_2)$. All the methods under the non-parametric bootstrap sampling scheme provide shorter confidence intervals than those obtained under the parametric scheme. The lower confidence bounds for percentile methods under both bootstrapping schemes shift more to the left compared to those obtained by other methods. The bootstrap BC_a confidence interval has the smallest expected length, capturing the asymmetry of the exact confidence intervals (CI).[28]

Table 4 Comparison of the bootstrap-t, bootstrap percentile (BP), Bias-corrected (BC) and BC_a confidence sets for $E(\hat{f}_2)\ (1-\alpha) = 0.90$

Method	Left	Right	Two-sided
	AV	**AV**	**EL**
Non-Parametric			
Bootstrap-t	60.61	67.36	6.75
BP	59.81	66.56	6.75
BC	60.65	67.39	6.74
BC_a	60.7	67.43	6.73
Parametric			
Bootstrap-t	60.66	67.56	6.9
BP	59.61	66.51	6.9
BC	60.45	67.24	6.81
BC_a	60.5	67.31	6.77

Conclusion

In this study parametric and non-parametric resampling methods are used to explore statistical properties of the similarity factor f_2. Under these two methods, 90% confidence intervals are constructed using different bootstrap approaches for the true expected value of $\hat{f}_2$.

For small sample sizes as in this study ($n_1 = 12$ for test drug, and $n_2 = 12$ for reference drug), nonparametric bootstrapping provides relatively smaller expected length of confidence intervals for the parameter f_2 compared to those obtained by the parametric method. However, the parametric bootstrap usually performs well over the non-parametric bootstrap method for small sample. In our study both methods provide similar CIs with no substantial advantage for considering one over the other. This may be due to the fact that the observed distribution of the reference drug was not normal. But we treated the observed distribution of the reference drug as normal to facilitate the parametric approach.

In addition to larger expected length confidence intervals, the parametric bootstrap method also provides less stable moment estimators of f_2 compared to the non-parametric bootstrapping methods. We note that BC_a performs the best in terms of producing smaller expected length among all the algorithms under both schemes. However, for small samples we recommend constructing bootstrap confidence intervals using non-parametric methods since these methods produce better results.

In this article, we showed that bootstrap is a powerful and effective means of setting approximate confidence intervals for the dissolution similarity measure f_2 using several computing algorithms. These results have policy implications for regulatory agencies such as FDA. Confidence intervals for the similarity factor f_2 provide more reliable prediction on the similarity of dissolution profiles of the test and the reference drugs.

Acknowledgements

None

Conflict of interests

Author declares that there is no conflict of interest.

References

1. Adams E. Evaluation of dissolution profiles using principal component analysis. *International Journal of Pharmaceutics*. 2001;212(1):41–53.
2. Huey Lin Ju. On the assessment of similarity of drug dissolution profiles–A simulation study. *Drug Information Journal*. 1997;31(4):1273–1289.
3. Vinod PS. In vitro dissolution profile comparison:–Statistics and analysis of the similarity factor, f2. *Pharmaceutical Research*. 1998;15(6):889–896.
4. Dawoodbhai S, Suryanarayan E, Woodruff C, et al. Optimization of tablet formulation containing talc. *Drug Develop Industrial Pharmacy*. 1991;17(10):1343–1371.
5. Pena Romero A, Caramella C, Ronchi M. Water uptake and force development in an optimized prolonged release formulation. *International Journal Pharmaceutics*. 1991;73:239–248.
6. Langenbucher F. Linearization of dissolution rate curves by Weilbull distribution. 1972.
7. Kervinen L, Yliruusi J. Modeling S–shaped dissolution curves. International Journal of Pharmaceuticals. 1993;92(1–3):115–122.
8. Tsong Y, Hammerstrom T, Lin KK, et al. The dissolution testing sampling acceptance rules. *Journal of Biopharmaceutical Statistics*. 1995;5(2):171–184.
9. Tsong Y, Hammerstrom T, Chen JJ. Multiple dissolution specification and acceptance sampling rule based on profile modeling and principal component analysis. *J Biopharm Stat*. 1997;7(3):423–439.
10. Moore JR, Flanner HH. Mathematical comparison of curves with an emphasis on dissolution profiles. *Pharma Technol*. 1996;20(6):64–74.
11. Chow SC, Liu JP. Design and analysis of bioavailability and bioequivalence studies. Chapman and Hall/CRC, 2008.
12. Tsong Y. Statistical assessment of mean differences between two dissolution datasets. Rockville: Presented at the 1995 Drug Information Association Workshop; 1995.
13. Tsong Y, Hammerstrom T, Sathe P, et al. Statistical assessment of mean differences between two dissolution data sets. *Drug Information Journal*. 1996;30(4):1105–1112.
14. DiCiccio TJ, Romano JP. A review of bootstrap confidence Intervals. *Journal of Royal Statistical Society*. Series B. 1988;50(3): 338–354.
15. Efron B. Bootstrap methods: Another look at the Jackknife. *Annals of Statistics*. 1979;7(1):1–26.
16. Chernick MR. Bootstrap Methods: A Guide for Practitioners and Researchers. 2nd edn. Wiley: Hoboken; 2008.
17. Davison AC, Hinkley DV. Bootstrap Methods and Their Applications. Cambridge University press, Cambridge: New York; 1997.
18. DiCiccio TJ, Efron B. Bootstrap confidence intervals. *Statistical Science*. 1996;11(3):189–212.
19. Efron B, Tibshirani R. An Introduction to the Bootstrap. , New York: Chapman and Hall; 1994.
20. Efron B. The Jackknife, the Bootstrap, and Other Resampling Plans. Philadelphia, Society for Industrial and Applied Mathematics, 1982.
21. Schenker N. Qualms about bootstrap confidence intervals. *Journal of American Statistical Association*. 1985;80(390):360–361.
22. Efron, B. (1981): *Nonparametric standard errors and confidence intervals*. Canadian Journal of Statistics, 9, 139–172.
23. Efron B. Better bootstrap confidence intervals. *Journal of the American Statistical Association*. 1981;82(39):171–185.
24. Bickel Pj, Freedman DA. Some asymptotics on the bootstrap. *Annals of Statistics*. 1981;9(6):1196–1217.
25. Hall P. *The Bootstrap and the Edgeworth Expansion*. New York: Springer–Verlag Hall; 1992.
26. Efron B. Nonparametric standard errors and confidence intervals. *Canadian Journal of Statistics*. 1981;9(2):139–172.
27. Carlos M Jarque, Anil K Bera. A test for normality of observations and regression residuals. *International Statistical Review*. 1987;55(2):163–172.
28. Tsong Y, Hammerstrom T. Statistical issues in drug quality control based on dissolution testing. Proceedings of the Biopharmaceutical Section of the American Statistical Association. 1992; 295–300.