dx.doi.org/10.14227/DT200113P13

PhEq_bootstrap: Open-Source Software for the Simulation of f_2 Distribution in Cases of Large Variability in Dissolution Profiles

e-mail: mfmendyk@cyf-kr.edu.pl

Aleksander Mendyk*, Adam Pacławski, Jakub Szlęk, and Renata Jachowicz

Department of Pharmaceutical Technology and Biopharmaceutics, Faculty of Pharmacy, Jagiellonian University–Medical College, Medyczna 9 St., 30-688 Kraków, Poland

ABSTRACT

PhEq_bootstrap is a free software tool that uses the similarity factor (f_2) to assess dissolution profile similarity in cases of large data variability. Its theoretical background is founded on bootstrapping, a statistical technique used to simulate the distribution of f_2 values based on the available sample. It allows both justification of profile similarity and prospective simulations for the establishment of the formulation development endpoint. The software is FOSS (free open-source software) and is available online (*1*).

INTRODUCTION

Moore and Flanner (*2*) proposed simple measures for the distance between the two curves representing dissolution profiles of two dosage forms, namely the difference (f_1) and similarity (f_2) factors. Because of its mathematical simplicity and lack of mechanistic assumptions, the similarity factor (f_2) is currently one of the most commonly used methods for the comparison of dissolution profiles (*3*). Although it is a simple method, f_2 is restricted to the following conditions (*4*):

- Minimum of three points in the profile.
- Twelve units for each reference and test product.
- No more than one point over 85%.
- RSD for dissolution points less than 10% (first point less than 20%).

The appropriate construction of the analytical protocol ensures the number and the range of the dissolution points applicable to the f_2 computation. However, the variability of the dissolution points is attributed mainly to the immanent characteristics of the dosage form (i.e., API solubility rate variations, coating non-uniformity, etc.). Thus, variability is usually beyond the control of the analyst, and its effect has to be minimized by appropriate statistical techniques to compare dissolution profiles with maximum reliability. Two major groups of techniques applicable here are based on covariance matrix and bootstrap.

Analysis of the covariance matrix is used in the Mahalanobis distance technique of the direct profile comparison. Since it is a multivariate technique of profile comparison, f_2 is not used here. The Mahalanobis distance technique suffers from the restrictions of the matrix inversion techniques, which are sensitive to the covariance matrix characteristics; thus, it is sometimes unable to handle real data. Moreover, interpretation of the Mahalonobis distance requires complex computations of the confidence intervals. Therefore, this approach is not as versatile and simple to interpret as f_2.

In contrast to the Mahalanobis distance technique, the bootstrap technique allows the use of f_2; however, it is also used not as a point estimator but as a confidence interval. The bootstrap technique is mathematically and algorithmically much simpler than the Mahalanobis distance technique. Moreover, it is easier to interpret based on the use of commonly known rules describing f_2 values and their meaning for the decision on the similarity of the analyzed profiles.

Based on the above considerations, the aim of this work was to create a computer program able to provide analysis of the f_2 confidence intervals with the bootstrap technique to help in the assessment of the similarity between dissolution profiles in cases of large variability in the dissolution data.

PhEq_Bootstrap FEATURES AND THEORETICAL BACKGROUND

PhEq_bootstrap is software that handles the above presented restrictions of f_2 concerning RSD of the dissolution points. The bootstrap technique is used to simulate f_2 distribution to assess the worst-case scenario as a lower confidence interval of the expected values of f_2. Computations of expected values and unbiased estimators of f_2 are based on the publication of Shah et al. (*5*), where mathematical proof of this concept is presented and discussed.

To carry out the computations, a new population of dissolution profiles was generated numerically by the bootstrap technique, where the new samples were the result of random sampling with replacement of the dissolution profiles. This procedure was performed for both the reference and test profiles. Two modes of the sampling proce-

*Corresponding author.

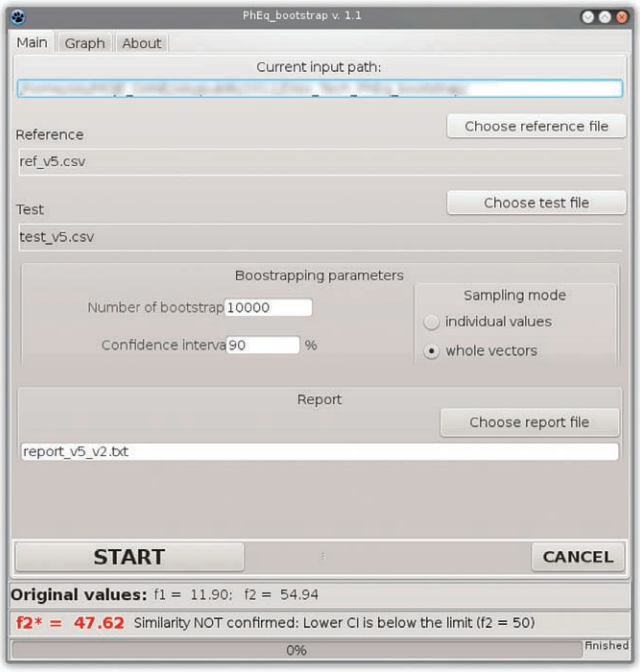

Figure 1. Main window of PhEq_bootstrap.

dure ("sampling mode") were implemented in this software (Figure 1). They were based on the definition of the sample, which consisted of whole profiles and individual points.

In Mode 1, the "new" dissolution profile is a simple replication of the profiles from the original data. The only difference is the set of profiles for each of the bootstrapped samples assigned to the reference and test drug. This means that a particular artificial sample for the reference or test drug contains the dissolution profiles from the randomly selected original profiles with possible replications (Figure 2). In the extreme case, one sample might contain 12 identical profiles from the same profile chosen by the system 12 times. In Mode 2, the dissolution profiles are created de novo, based on the sampling of the individual points. For each time point available in the original data, the amounts of released drug are treated as a pool from which a randomly chosen value is used for the new, numerically created dissolution profile (Figure 3). Both modes are carried out separately for the test and reference drug with sampling based on flat (uniform) distribution. The number of the artificial dissolution profiles (the bootstrap number) is a subject of the individual optimization. It means that the user should perform a few runs of the software consecutively, with an increasing number of bootstraps (i.e., 500, 1000, 2000, etc.) and observe the results. Usually 1000 bootstraps or more is enough to stabilize the results

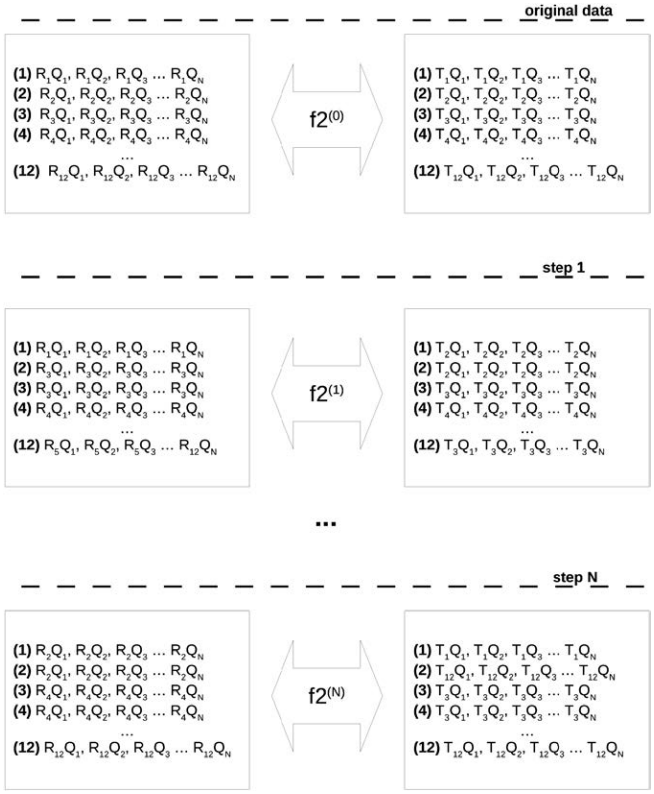

Figure 2. Mode 1 of the bootstrapping algorithm implemented in PhEq_bootstrap.

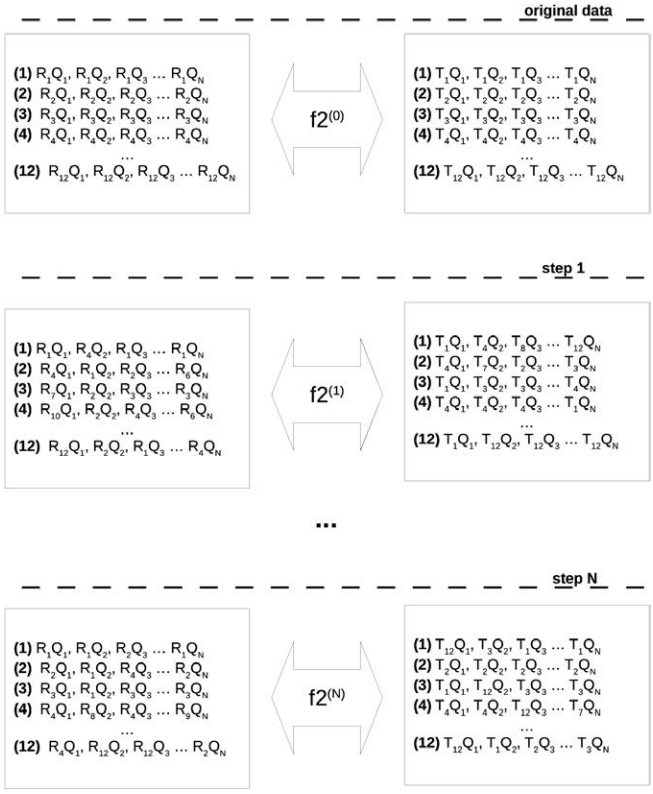

Figure 3. Mode 2 of the bootstrapping algorithm implemented in PhEq_bootstrap.

(*5*), thus they might be regarded as final. The above rule is empirical and is therefore not valid in all cases; the software default is 5000 and may be increased if needed. Another default value is the confidence interval (CI) set to 90%, which might be changed with appropriate discussion. The considerations presented above are very important for the validation of the software. The software is based on the probabilistic technique (bootstrapping) where the (pseudo) random numbers are the base of the confidence interval computations. This implies that running the analysis several times would lead to slightly different results. The more bootstraps set for a single run, the smaller would be the differences between the results of consecutive runs, yet at a certain level of accuracy they would always exist.

An interpretation of the results is based on the common rule of dissolution profile similarity: $f_2 > 50$. However, in this case when confidence intervals are known, the above rule is applied to the lower CI of expected value of f_2 (*5*), which makes this reasoning the "worst-case scenario" approach. The characteristics of the original dissolution results are the basis for the estimation of the "true" value of f_2 expressed as the CI simulated with the bootstrapping technique. In the end, one selects the least favorable value of the CI; since f_2 should be greater than 50, it is the lower CI value. Such interpretation is performed automatically by the software, and the results are displayed immediately at the end of the computations. This is a simple go/no-go decision. However, other valuable information derived from the computations of the bootstrapped samples is the difference between the average value of f_2 and its corresponding lower confidence interval value. It may provide a rough estimation of an endpoint of the average f_2 value to be achieved to maintain similarity of the profiles in cases when the bootstrapping analysis reveals a lack of similarity (see the case studies below).

Operation of the software is very simple and requires only a few steps (Figure 1):

- Reference and test samples must be defined by loading tab-delimited txt/csv files by pressing the button "Choose {reference, test} file."
- The report file should be also defined: "Choose report file."
- Optionally, the bootstrapping parameters might be changed from the defaults.
- The button "Start computations" runs the whole procedure.

The progress is monitored and displayed at the bottom of the main window in the form of a progress bar (Figure 1). It must be emphasized that these are iterative operations; therefore, it might take time to get the results depending on the number of bootstraps and the computer speed. Nevertheless, for default parameters it usually takes less than a minute to get the results. Simulations performed with an unrealistic 500,000 bootstraps lasted 70 min at most on the average computer. In any case, if computations have to be stopped, please use the "Cancel" button. However, all the progress would be lost, and the next time the computations would be resumed from the beginning. For reference, the software also computes f_1 for original and bootstrapped profiles.

The criterion that no more than a single point is above $Q = 85\%$ is applied automatically, and therefore the software displays the actual number of points used for the computation of f_1/f_2 in the report file. The current version of *PhEq_boostrap 1.1* is available as Open-Source software from the SourceForge server (*1*).

CASE STUDIES

Case studies 1 and 2 were prepared to demonstrate how similarity of the dissolution profiles is influenced by the results variability. The data are simulated and are available from the SourceForge server with the *PhEq_boostrap* package (*1*). Various datasets were prepared for the discussion:

- *reference_set.csv* – for both cases
- *test_set_1.csv* – for case 1
- *test_set_2.csv* – for case 2

Case Study 1

Based on the distance between the average reference and test profiles, it looks like both products might be considered similar (Figure 4). Their corresponding f_2 value seems to support this claim: $f_2 = 54.94$. However, data variability discourages the use of the direct f_2 computation; for certain time points, RSDs are greater than allowed (Table 1). Analysis with *PhEq_bootstrap* leads to contradic-

Table 1. Average Dissolution Profiles for Case Study 1 with Variability Assessment

Time (min)	Q	SD	RSD (%)
	Reference		
3	15.34	3.23	21.07
5	36.57	4.07	11.13
7	50.49	5.95	11.78
10	68.47	6.50	9.49
15	77.63	6.39	8.24
30	90.23	7.32	8.12
	Test		
3	8.01	3.42	42.72
5	21.80	3.55	16.30
7	43.06	4.28	9.94
10	65.83	6.56	9.97
15	75.50	8.75	11.59
30	84.23	10.33	12.27

SD: standard deviation
RSD: relative standard deviation

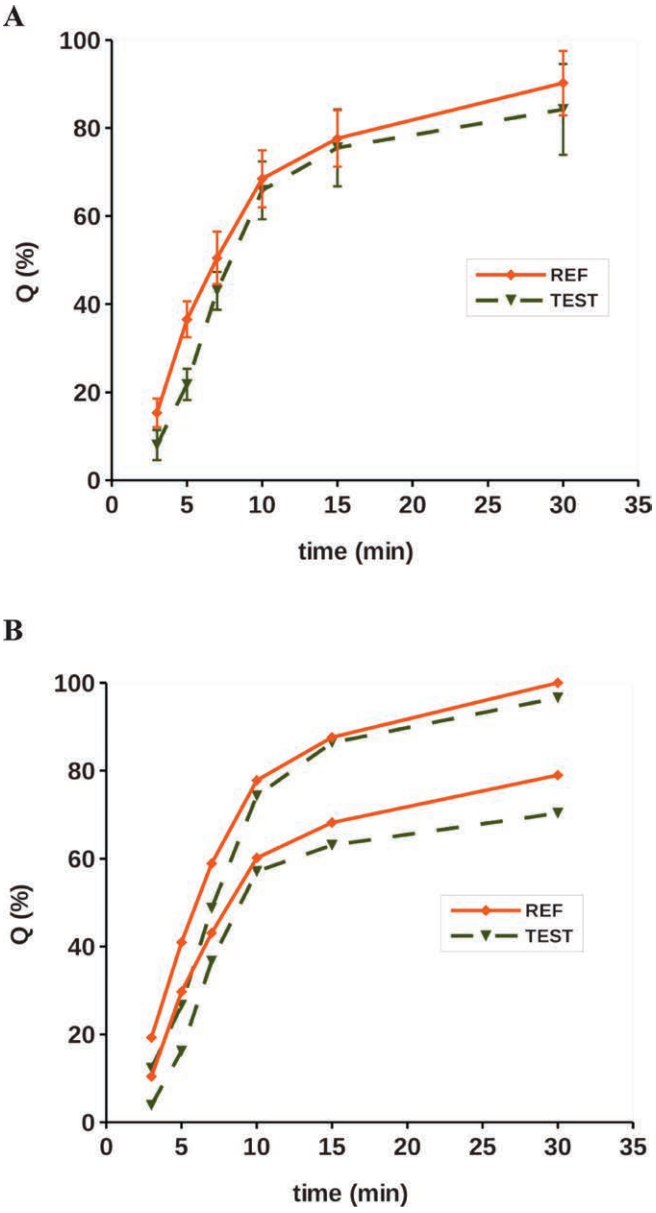

Figure 4. Dissolution profiles of test vs. reference for Case 1. (A) Average profiles with bars representing ±SD; (B) maximum/minimum range of the dissolution profiles of test and reference product.

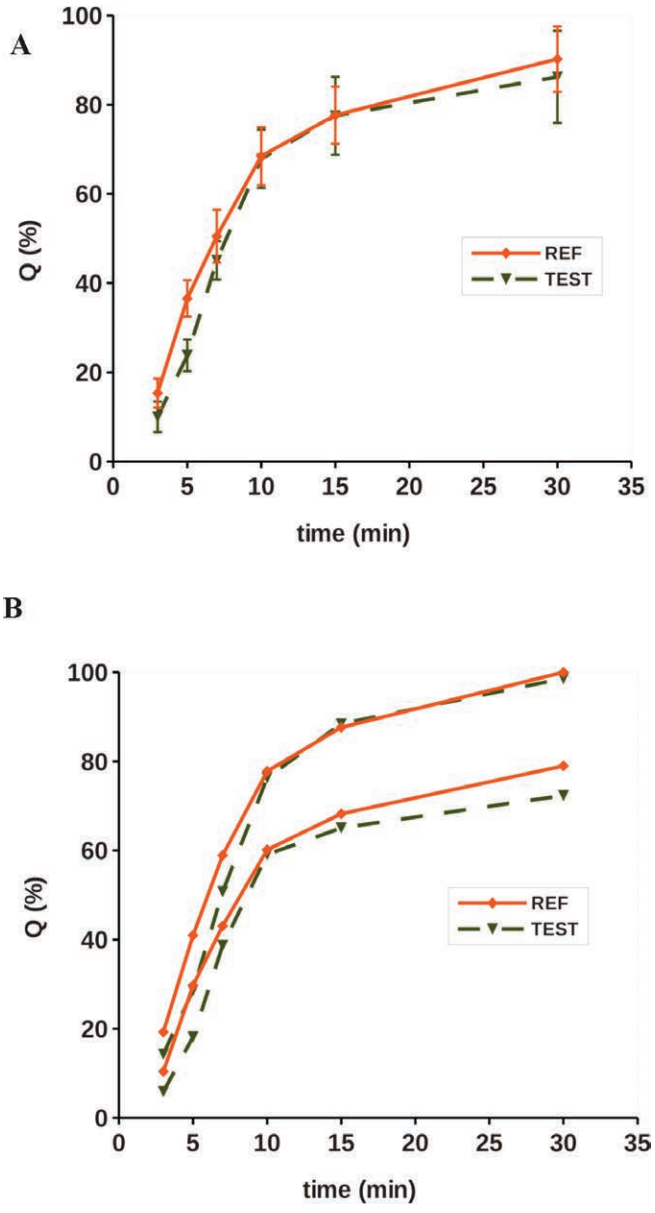

Figure 5. Dissolution profiles of test vs. reference for Case 2. (A) Average profiles with bars representing ±SD; (B) maximum/minimum range of the dissolution profiles of test and reference product.

tory conclusions about profile similarity. The lower CI of the expected value of f_2 (f_2*) is 47.62, which is below the criterion of $f_2 > 50$. Based on the above results, the profiles cannot be confirmed as similar.

Case Study 2

In the first case study, because the unsatisfactory f_2* values did not confirm the similarity of the profiles, the question becomes, what is the minimum f_2 value computed for the average profiles that would confirm the similarity of profiles with such variability? To answer this question, the test profiles were simply shifted to move up their average values by the factor $Q = 2\%$ (Figure 5).

In consequence, the profiles were closer to each other, but the variability characteristics were unchanged. Accordingly, the resulting point estimator of $f_2 = 59.80$ is more favorable, and the analysis with *PhEq_bootstrap* results in f_2* = 51.41. The profiles were confirmed as being similar despite their large variability. This case study shows the possible use of *PhEq_bootstrap* as a simulation tool for identification of the formulation endpoint for generic products in particular. The endpoint is understood as the minimum f_2 value to guarantee its similarity to the reference, despite the large variability of the dissolution profiles. In this example, it might be concluded that such an f_2 value should be no less than 60.

SUMMARY

PhEq_bootstrap is a free software tool to help with similarity factor (f_2) calculation to assess dissolution profile similarity in cases of a large variability of the dissolution profiles. Its theoretical foundations include the bootstrapping technique and analysis of the expected value of f_2. It allows both the justification of the profile similarity and prospective simulations for the establishment of the formulation development endpoint. *PhEq_bootstrap* is available under GPLv3 license (*6*) from sourceforge.net (*1*), therefore it is free to use for private and commercial purposes.

ACKNOWLEDGMENTS

This software was released as open-source software and was created with the sole use of the open-source tool Lazarus (*7*) working under openSUSE Linux environment (*8*). Its creation was not supported financially by any funds or grants. The only support received was the knowledge derived from the outstanding work of Shah et al. (*5*) and the great commitment of the open-source community providing software development tools and operating systems.

REFERENCES

1. SourceForge *PhEq_bootstrap* Web site. http://sourceforge.net/projects/pheqbootstrap (accessed Jan 13, 2013).
2. Moore, J. W.; Flanner, H. H. Mathematical Comparison of Dissolution Profiles. *Pharm. Technol.* **1996**, *20* (6), 64–74.
3. Shah, V. P.; Tsong, Y.; Sathe, P.; Williams, R. L. Dissolution Profile Comparison Using Similarity Factor, f_2. *Dissolution Technol.* **1999**, *6* (3), 15.
4. *Dissolution Testing of Immediate Release Solid Oral Dosage Forms*; Guidance for Industry; U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), U.S. Government Printing Office: Washington, DC, 1997. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070237.pdf (accessed Jan 13, 2013).
5. Shah, V. P.; Tsong, Y.; Sathe, P.; Liu, J. P. In Vitro Dissolution Profile Comparison—Statistics and Analysis of the Similarity Factor, f_2. *Pharm. Res.* **1998**, *15* (6), 889–896.
6. GNU Operating System Home Page. http://www.gnu.org/licenses/gpl.html (accessed Jan 13, 2013).
7. Lazarus Home Page. http://www.lazarus.freepascal.org (accessed Jan 13, 2013).
8. openSUSE Linux Operating System Home Page. http://www.opensuse.org (accessed Jan 13, 2013).