Comparison of Dissolution Profiles with Heterogeneous Variability between Test and Reference Batches

Xiu Huang[1], Patrick Marroum[2], Weihan Zhao[1]

[1] Discovery and Exploratory Statistics, AbbVie Inc. ; [2] Clinical Pharmacology and Pharmacometrics, AbbVie Inc.

1 Introduction

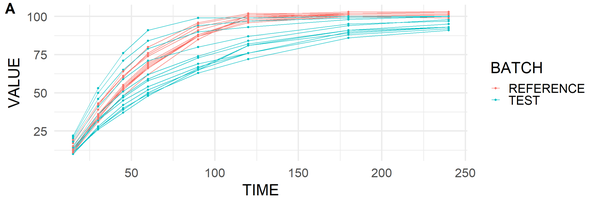
1.1 Background

- *In vitro* dissolution testing is critical for drug quality control and assess similarity of release characteristics in granting *in vivo* bioavailability/ bioequivalence waivers.
- For highly variable dissolution profiles, multivariate model-independent procedures are recommended by FDA and EMA guidelines.
- However, these approaches have been developed with the underlying assumption of **`homogeneous test-reference variances`**.

1.2 Objectives

Motivated by an in house dissolution dataset where the test and reference batches have **`unequal variabilities`** (**Figure 1.1**), we'd like to study the performances of existing approaches, propose novel dissolution testing methodologies to account for the variance heterogeneity, and revisit the definition of **`equivalence`**.

A



B

RSD	15min	30min	45min	60min	90min	120min	180min	240min
REFERENCE	19.4%	10.68%	8.68%	7.41%	4.03%	1.85%	0.88%	1.07%
TEST	29.92%	26.24%	25.32%	23.31%	16.56%	10.45%	5.08%	3.64%

Figure 1.1: Dissolution profiles from two batches of an in house dissolution dataset. Dissolution testing is conducted to assess bioequivalence. It is obvious from the visualization of the data and the table of the relative standard deviation (RSD) that the two batches are of unequal variances.

2 Methods

2.1 Dissolution Data

- Dissolution data set of the reference formulation from the motivating data set (**Figure 1.1**, red) was used as the common reference set.
- Dissolution data set of the test formulation was simulated based on the motivating data set (**Figure 1.1**) to mimic its statistical property but with varying relative standard deviation (RSD) levels.

2.2 Commonly Used Methodologies

- **Similarity factor** f_2 (Shah et al. 1998): for highly variable dissolution data when RSD is more than 20% or 10% at early or later time point respectively, f_2 **`does not apply`**.
- **Bootstrapped** f_2 (Islam 2018): the bias corrected and accelerated (BCA) confidence interval is in general more commonly used.
- **Multivariate Statistical Distance (MSD)** (Tsong et al. 1996): adopted the Mahalanobis distance as the multivariate distance measure; **`homoscedasticity`** for test and reference profiles is assumed. Global 10% difference is used to calculate the equivalence margin.
- **T2EQ** (Hoffelder 2018): multivariate model-independent procedure using the Wellek's T^2-test, which was claimed to present the best compromise between type I error and power (Suarez-Sharp et al. 2020); **`homoscedasticity`** for test and reference profiles is assumed. Global 10% difference is used to calculate the equivalence margin.

2.3 Exploring Novel Methodologies

- **Generalized Mahalanobis Distance (GMD)** (Hoffelder, Gössl, and Wellek 2015): adjusted Mahalanobis distance measure which involves the **`entire covariance structure`** of the test and reference data. A data driven equivalence margin is calculated with 10% global difference.
- **Modified T2EQ (MT2EQ)**: built on top of the T2EQ framework, utilizing the **`Krishnamoorthy and Yu testing statistics`** for the multivariate Behrens–Fisher problem (Krishnamoorthy and Yu 2004). A data driven equivalence margin is calculated with 10% global difference.

3 Results

3.1 Motivating Data (Figure 1.1) with Conflicting Results

- The f_2 value calculated was 52.05, **`passing`** the threshold of 50. However, the RSDs for test batch were more than 20% and 10% at early and later time points respectively, didn't qualify for the f_2 application.
- The **bootstrapped** f_2 with the 90% BCA CI calculated was 52.1 (49.3,58.2). The lower limit was less than 50, therefore **`did not pass`** the equivalence test.
- The **MSD** calculated with the 90% confidence interval was 6.921 (4.885,8.956). The upper bound was less than the 10% global difference margin (10.15), indicating the two batches **`pass`** the equivalence test.
- The **T2EQ**, **GMD**, **MT2EQ** test were **`passed`** with P-value 6.4e-05, 8.5e-05, and 5.1e-13 respectively.

3.2 Simulated Data Set

- **Simulation 1**: Test profiles were simulated from a **`multivariate normal distribution`** with means parallel to the reference profile, and with RSD from the same level up to ~10 times of the reference RSD. Heterogeneous AR(1) covariance ($\rho = 0.9$) was assumed.

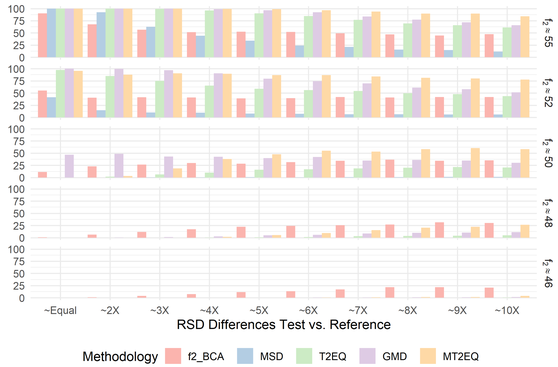


Figure 3.1: Percentage of experiments that pass the equivalence test for simulation 1. 1000 test profiles were simulated for each of the variance level.

- **Simulation 2**: Test profiles were simulated from a **`multivariate normal distribution`** with means parallel to the original test profile (therefore not parallel to the reference profile), and with RSD from same level up to ~10 times of the reference RSD. Heterogeneous AR(1) covariance ($\rho = 0.9$) was assumed.

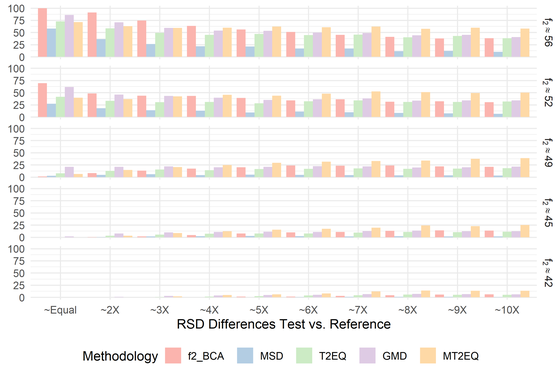


Figure 3.2: Percentage of experiments that pass the equivalence test for simulation 2. 1000 test profiles were simulated for each of the variance level.

3.3 Results Discussion

- When analyzing the motivating data, different conclusions could be drawn using different methodologies.
- It remains a question as what is the desired outcome when comparing the test and reference batches when they are of different variances. Does the **`"equivalence"`** refer to the profile mean being equivalent, or should the profile variance be equivalent as well? Only when this question is answered will a proper equivalence testing procedures be designed and the performance evaluations be justified.
- If only the mean equivalence is relevant, the newly explored **`GMD`** and **`MT2EQ`** in general have higher power than the other methods, although they show higher but still limited type I error when the variance differences increase.
- When the test and reference profiles are parallel with each other, multivariate distance based methods in general will have more power compared with bootstrap f_2.

4 Next Steps

- FDA has previously guided that to declare equivalence, one needs to compare the variances of the dissolution profiles first, and **`justify variance difference`** if there's any before proceeding to statistical testing step. It would be an interesting topic to explore the possibility of factoring the two-step dissolution testing into an one-step procedure and test its practicality in different settings.
- How to properly specify the **`equivalence margin`** for multivariate distance based methods remains challenging. For the two newly proposed methods and other potential multivariate methodologies, further exploration of different equivalence margin settings is needed.

References

Hoffelder, T. 2018. "Comparison of Dissolution Profiles: A Statistician's Perspective." *Ther Innov Regul Sci* 52 (4): 423–29.

Hoffelder, Thomas, Rüdiger Gössl, and Stefan Wellek. 2015. "Multivariate Equivalence Tests for Use in Pharmaceutical Development." *Journal of Biopharmaceutical Statistics* 25 (3): 417–37.

Islam, Mohammad. 2018. "Bootstrap Confidence Intervals for Dissolution Similarity Factor F2." *Biometrics & Biostatistics International Journal* 7 (September). https://doi.org/10.15406/bbij.2018.07.00237.

Krishnamoorthy, K, and Jianqi Yu. 2004. "Modified Nel and van Der Merwe Test for the Multivariate Behrens–Fisher Problem." *Statistics & Probability Letters* 66 (2): 161–69.

Shah, V. P., Y. Tsong, P. Sathe, and J. P. Liu. 1998. "In vitro dissolution profile comparison-statistics and analysis of the similarity factor, f2." *Pharm. Res.* 15 (6): 889–96.

Suarez-Sharp, S., A. Abend, T. Hoffelder, D. Leblond, P. Delvadia, E. Kovacs, and D. A. Diaz. 2020. "In Vitro Dissolution Profiles Similarity Assessment in Support of Drug Product Quality: What, How, When-Workshop Summary Report." *AAPS J* 22 (4): 74.

Tsong, Yi, Thomas Hammerstrom, Pradeep Sathe, and Vinod P. Shah. 1996. "Statistical Assessment of Mean Differences Between Two Dissolution Data Sets." *Drug Information Journal* 30 (4): 1105–12. https://doi.org/10.1177/009286159603000427.

This poster was sponsored by AbbVie. AbbVie contributed to the design, research, and interpretation of data, writing, reviewing, and approving the publication. All authors are employees of AbbVie Inc. and may own AbbVie stock.