abbvie

Rational Statistical Analysis Practice in Dissolution Profile Comparison for Product Quality Assessment of Similarity through Real Case Studies: Industry Perspective

May 21, 2019

Yanbing Zheng, Jian-Hwa Han, James Reynolds, Mark Johnson, Karin Rosenblatt, Tzuchi R Ju, Yi Gao, Bei Chen, Hesham Fahmy This presentation was sponsored by AbbVie. AbbVie contributed to the design, research, and interpretation of data, writing, reviewing, and approving the publication. The authors are employees of AbbVie, Inc.

Outline

- Model independent statistical methods
- Simulation studies
- Decision tree
- R Shiny tool
- Case studies



f_2 Rules (FDA 1997 Guidance)

- N=12 of (i) Reference (or prechange) and (ii) Test (or postchange) products
- Use the Mean values only for calculation
- Model Independent Method most suitable for dissolution profile comparison when three to four or more dissolution time points are available
 - Same time points (minimally 3 times points)
 - Only one measurement should be considered after 85% dissolution of both the products
 - %RSD NMT 20% at early points (e.g. 10 minutes);
 NMT 10% for all other points

What if f_2 assumptions are not satisfied?

- It is critical to identify a right tool/method in order to make meaningful assessment for product quality
 - Model independent statistical methods
 - \circ f_2 bootstrap (Shah, et al. 1998)
 - Tsong's MSD method (Tsong, et al. 1996)
 - SK method (Saranadasa and Krishnamoorthy 2005)
 - Saranadasa's Hotelling's T² based method (Saranadasa 2001)
 - Intersection union test (Berger and Hsu 1996)
- Simulation studies were performed to evaluate the power and type I error of different approaches.
- More than 250 cases were used for the establishment of decision tree and assessment.

5

Model Independent Statistical Methods

Methods are based on some function of the distance between the profiles at each time point

- f_2 Euclidean distance (pythagorean theorem) based on equal weights (1/p)
- Tsong's MSD and Hotelling's T²– Euclidean distance weighted by standard deviations and correlations
- SK common distance weighted by complex function of standard deviations and correlations
- Intersection Union Test maximum distance weighted by standard deviations

6

Statistical Methods for Dissolution Profile Comparison

Methods	Pros	Cons	Comments
Similarity factor f_2	 Simple Common acceptable cutoff: 50 	 Uses only the mean profile Loses applicability when variability increases Lack of type I error control Unknown statistical distribution 	 FDA requirements: %CV <=20% at the earlier time points and <=10% at other time points.
<i>f</i> ₂ bootstrap	 Considers profile mean and variation Common acceptable cutoff: 50 	Could be conservative	 Recommended when f₂ usage requirements on variation are exceeded. Strong regulatory connection.
Tsong's Multivariate statistical distance (MSD) method	 Considers profile mean and variation Real case studies suggest good statistical power of claiming similarity and type l error control. 	Cutoff is random and data dependent	 No common acceptable cutoff. Strong regulatory connection.

Methods	Pros	Cons	Comments
Saranadasa and Krishnamoorthy's (SK) method	 Considers profile mean and variation Cutoff 10% approximately corresponds to f₂ 50 	 Assumes parallelism of the two dissolution profiles Liberal. 	 The assumption is usually not satisfied in practice.
Sarandasa's Hotelling T ² -based method	 Considers profile mean and variation Cutoff value 6% was proposed. 	 Assumes parallelism of the two dissolution profiles 	 The assumption is usually not satisfied in practice.
Intersection-Union Test	 Considers profile mean and variation Be able to identify the time-point(s) that does not show similarity 	 Time points are considered independently Very conservative 	Too conservative
Model-dependent approaches	 Measurements can be taken at different time points for reference and test batches. 	 Model selection Spacing of time points may limit curve/model choices Cutoff selection 	 Appropriate when dissolution curves are sampled at many time points. Hard to have a common acceptable cutoff.

Simulation Study

- Mean for test profile =(35, 45, 70, 85) and compound symmetry covariance structure with correlation=0.5.
- Assume equal covariance matrices.
- Assume parallelism between reference and test dissolution profiles (δ: constant difference over time points between two profiles)
- Consider various variability
 - RSD%=(5.7, 4.4, 2.9, 2.4)% for test profile
 - RSD% = (14.3, 11.1, 7.1, 5.8)% for test profile
 - RSD% = (28.6, 22.2, 7.1, 5.8)% for test profile
- For each variability and δ , 1000 simulated data sets were generated to assess probability of claiming equivalence.

9

Simulation Study – RSD%=(5.7, 4.4, 2.9, 2.4)% for test profile



- All methods have high power to claim similarity for small δ
- Bootstrapped f2 and SK give probability of claiming equivalence close to 5% when δ=10%

10

Simulation Study – RSD% = (14.3, 11.1, 7.1, 5.8)% for test profile



Simulation Study – RSD% = (28.6, 22.2, 7.1, 5.8)% for test profile



- f2 assumptions are violated.
- Comparing to SK, f₂
 bootstrap and MSD method are relatively conservative for highly variable cases.

Simulation Study – Method Comparison

• Assume equal covariance matrices and RSD% = (28.6, 22.2, 7.1, 5.8)% for test profile

Similarity passing rate	Mean Diff= (28, 22, 10, 5) f2=36.7	Mean Diff= (18, 13, 8, 5) f2=45.8	Mean Diff= (12, 10, 9, 5) f2=51.3	Mean Diff= (10, 10, 3, 3) f2=56.4	Mean Diff= (5, 4, 3, 3) f2=70.1
f2	0.001	0.203	0.586	0.822	0.982
Bootstrapped f2	0	0.014	0.094	0.257	0.728
MSD	0.005	0.054	0.073	0.373	0.623
f2>=50 & (Bootstrapped) f2 or MSD)	0.001	0.041	0.137	0.445	0.804
SK	0.410	0.625	0.529	0.974	0.978
IUT	0	0	0.006	0.024	0.162
Good power and type I error Caution!					

Summary/Remarks

- IUT is very conservative and has very low power to claim similarity.
- SK method has good power to detect similarity and control of type I error when the two dissolution profiles are parallel. But when the underlying assumption of parallelism fails, SK method could be too liberal with high type I error (pass similarity when dissimilar).
- Comparing to SK, f2 bootstrap and MSD method are relatively conservative for highly variable cases.
- MSD is inconsistent in its result comparing to bootstrapped f2. MSD method is likely to be less discriminating and sensitive in some scenarios (e.g. Paixão, et al. 2017 and Mangas-Sanjuan, et al. 2016). But on the other hand, MSD method can also have higher power to detect similarity in some scenarios when the two profiles are similar.
- f2 is a conservatively biased estimator. Although f2 and MSD are testing different hypotheses, comparisons may fail bootstrap and pass MSD in part because of the conservative bias of f2.

Decision Tree



Three Methods are utilized in this practice: f2; f2 Bootstrapping; MSD (Tsong's Method)

Scenario 1: f2≥50 & Variability met (Pass/ --- / ---) Scenario 2: f2<50 (Fail/ --- / ---)

NOT met Variability requirements: Scenario 3: (Pass/ Pass / ---) Confirmed by Bootstrapping! Scenario 4: (Pass/ Fail / Fail) → Cannot confirm similarity Scenario 5: (Pass/ Fail / Pass) Confirmed by MSD method!

Basic Concepts for the Decision Tree

Three methods in series for analysis based on the f2 criteria

- f2 Calculation
- f2 Bootstrapping (more conservative than f2)
- Tsong's MSD method (additional checking for borderline cases)



R Shiny Web Application Tool established

Dissolution Profiles Comparison

Choose Ref lot to upload
Browse Reference.csv
Upload complete
☑ Header
Choose Test lot to upload
Browse Test.csv
Upload complete
☑ Header
Time Unit
minutes -
All time points (eg, 10,20,30)
10,15,30,45,60,90
Number of time points for comparison (>= 3)
4
Plot Title
Dissolution Profile Comparison
Submit
Senerate report Sur Instruction



Methods	Statistics.Value	Reference.Value	Pass.Or.Not	Recommend
f2	60.22	>=50	PASS	0
f2 bootstrap	50.66	>=50	PASS	1
Tsong MSD	3.43	<= 11.55	PASS	0

The study belongs to Scenario #3. Similarity is confirmed.

Note: If similarity is confirmed, then the recommended method for similarity assessment will be highlighted yellow with value 1 in the Recommend column.

Warning: The assumptions of f2 method are not satisfied. The CV for the later time points exceeds 10%.

Case Studies Summary

Prototype Formulations (Research Data, Total of 250 cases) – Confirmation rate is higher for lower release cases

	Pass!	Pass!	???	Fail!	Fail!		Confirmed
Categories	1	P/P/-	P/F/P	P/F/F	F/-/-	Total:	by Decision Tree*
>80% LA	11	66	16	9	28	130	87.7%
50% ~80%	1	4	2	0	3	10	80.0%
20% ~50%	0	30	3	1	6	40	92.5%
<20%	0	70	0	0	0	70	100.0%
	12	170	21	10	37	250	91.6%
	* The cases	can be clear	ly identified	as either Pa	iss or Fail by	the decisio	on tree.

➔ More than 85% of cases can be confirmed by this decision tree practice

Bootstrapping Performance – Some Cases May Desire the MSD Analysis



EMA/810713/2017 – Q&A on Mahalanobis distance

Question and answer on the adequacy of the Mahalanobis distance to assess the comparability of drug dissolution profiles

Draft agreed by Biostatistics Working Party	June 2018
Adopted by CHMP	26 July 2018

Keywords	Bioequivalence, dissolution profiles, f2, Mahalanobis distance, biowaiver
----------	---

"Based on these considerations, the MD metric cannot be supported as a preferred methodological approach to decide upon similar dissolution,"

Example of Scenario (F/F/P) – EMA's Concern on MD/MSD method



This case fails the similarity analysis according to our decision tree since f2 <50. MSD method may not be reliable if used alone.

Example of Scenario (P/F/P) – Very Similar Profiles Fail Bootstrapping due to high variability



Research data:

Reference and Test samples are both variable \rightarrow Similar! (MSD method is OK.) (May require N=12 to confirm!)

Example of Scenario (P/F/P) – Data Reliability?



The Reference samples are very variable → Data is not reliable! (May require New sample or Re-Test!)

Low Release Case of Scenario (P/P/-):



The Reference samples are very consistent (low variability), but the Test samples showed different behavior → Not Similar!

Low Release Case of Scenario (P/F/P):



The overall release is low and individual samples are not overlapping → Not Similar!

Dissolution Profiles Comparison – Factors to Consider & The Intention of Decision-Tree Practice

Factors to Consider \rightarrow Obtain Reliable Data for Comparison

- Sample Formulation Design, i.e. IR vs. ER
- Method
 - Hydrodynamics:
 - Apparatus Types
 - RPM/DPM/Flow Rate
 - Medium pH Physiological pH ranges

Statistical Tools (Methods) help us to understand the situation
 →Identify the Root Cause & Fix it
 Intention →Improve and Assure Product Quality
 (NOT just trying to Pass Similarity Analysis)

26

Attempting to let the science inform decision making.

NOT trying to pass products which are "dissimilar".

Nor are we wanting to fail products which are "similar".

The decision tree is not intended for use with every profile comparison situation. Check the science and the assumptions on the use of the statistical methods first.

If f2 < 50, then no need to test further as this implies there is more than 10% difference between the means of the test and reference

Some "similar" cases which fail bootstrap pass MSD

f2 is a conservatively biased statistic

References

- Berger, R. L. and Hsu, J. C. (1996). Bioequivalence trials, intersection-union tests and equivalence confidence sets. *Statistical Science* 11, 283-319.
- FDA's Guidance for Industry, Dissolution Testing of Immediate Release Solid Oral Dosage Forms. August 1997.
- Mangas-Sanjuan, V., Colon-Useche, S., Gonzalez-Alvarez, I., Bermejo, M. and Garcia-Arieta, A. (2016). Assessment of the regulatory methods for the comparison of highly variable dissolution profiles. *The AAPS Journal*. DOI: 10.1208/s12248-016-9971-5
- Paixão, P., Gouveia, L. F., Silva, N. and Morais J. A. G. (2017). Evaluation of dissolution profile similarity Comparison between f2, the multivariate statistical distance and the f2 bootstrapping methods. *European Journal of Pharmaceutics and Biopharmaceutics* 112, 67-74.
- Saranadasa, H. (2001). Defining the similarity of dissolution profiles through Hotelling's T2 Statistic. *Pharmaceutical Technology* 25, 46-54.
- Saranadasa, H., and Krishnamoorthy, K. (2005). A multivariate test for similarity of two dissolution profiles. *Journal of Biopharmaceutical Statistics* 15, 265-278.
- Shah, V. P., Tsong, Y., Sathe, P., and Liu, J.-P. (1998). In vitro dissolution profile comparison Statistics and analysis of the similarity factor, f₂. *Pharmaceutical Research* 15, 889-896.
- Tsong Y, Sathe PM, and Shah VP (2003) In vitro dissolution profile comparison. In *Encyclopedia of* Biopharmaceutical Statistics, Second Edition, Chow, S.-C., Ed.; CRC Press: Boca Raton, FL, pp 456-462.
- Tsong, Y., Hammerstrome, T., Sathe, P., and Shah, V. P. (1996). Statistical assessment of mean differences between two dissolution data sets. *Drug Information Journal* 30, 1105-1112.
- Wu, Sutan (2014). Current statistical issues in dissolution profile comparisons. 2014 Midwest Biopharmaceutical Statistics Workshop.
- Zheng, Y. and Zhang, L. (2018). Dissolution profiles comparison: Model independent approaches. In Encyclopedia of Biopharmaceutical Statistics, Fourth Edition, Chow, S.-C., Ed.; CRC Press: Boca Raton, FL, . DOI: 10.1081/E-EBS3-140000011

obbvie