Alternatives to f2 Testing for Dissolution Similarity – f2 Bootstrapping and Multivariate Statistical Distance (MSD) Method

SBIA 2020: Advancing Innovative Science in Generic Drug Development Workshop

Session 4: Practical Considerations in the Study Design and Data Evaluation Recommended in Product Specific Guidances

Topic 1: Oral Products

Xiajing (Jean) Gong, PhD.
Division of Quantitative Methods and Modeling
Office of Research and Standards
OGD | CDER | U.S. FDA
September 30, 2020

Disclaimer

This presentation reflects the views of the author and should not be considered to represent FDA's views or policies.

Learning Objectives

- Understand and describe the frequently used statistical approaches for dissolution profile similarity assessment in bioequivalence (BE) determination
- Understand how to apply f2 bootstrapping and MSD method for dissolution similarity testing via case studies

Dissolution Profile Similarity

- In vitro dissolution profile comparison is important for the evaluation of generic drug products
  - e.g., biowaiver for other strengths based on BE of biostrength, BCS-based biowaiver
- Dissolution profiles may be considered similar by virtue of overall profile similarity and similarity at every dissolution sample time point
- FDA guidance has recommended several statistical approaches for comparing dissolution profiles
  - Model independent: similarity factor f2, multivariate confidence region procedure
  - Model dependent approaches
- f2 testing has been considered the most widely applicable method for assessing similarity between two dissolution profiles

Similarity Factor f2

$$f_2 = 50 \bullet \log \{[1+(1/n)\sum_{t=1}^{n}(R_t - T_t)^2]^{-0.5} \bullet 100\}$$

N is the number of time points. R_t and T_t are the mean % API dissolved at time point t for Reference batch and Test batch, respectively

- 12 units (each in own dissolution vessel) for each product
- Minimum of 3 time points (zero excluded)
- Only one measurement should be considered after 85% dissolution of both the Test and Reference products
- Dissolution measurements should be made under same conditions and the dissolution profiles should have the same time points
- Requirements on variability

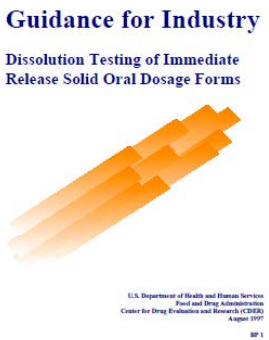
- f2 = 100 would mean the mean difference at each time point is 0.
- If f2 ≥ 50 the two profiles are considered “bioequivalent” or “similar”

Alternative Approaches to f2 Testing

- Requirement on **variability** for f2 calculation
  - Coefficient of variation (%CV) should not be more than 20% at the earlier time points (e.g., 15 minutes)
  - %CV should not be more than 10% at other time points

- Current regulatory practice for highly variable dissolution data
  - Model independent statistical method
    - f2 bootstrapping
    - Multivariate statistical distance (MSD) method

i.e., B. Model Independent Multivariate Confidence Region Procedure



f2 Bootstrapping

- Generate N (e.g., N=10,000) bootstrap samples by resampling with replacement from dissolution data for the test and reference products
- Estimate f2 for each of the N bootstrap samples
- Calculate the bootstrapped f2 mean which represents the average of N f2 values
- Calculate 90% confidence interval of the bootstrapped f2 using the **percentile method** (bias-corrected and accelerated bootstrap (BCa) approach may be applied)

Multivariate Statistical Distance (MSD) Method

Distance measure: Mahalanobis distance

$$D_M = \sqrt{(R_t - T_t)'\left(\Sigma_{pooled}\right)^{-1}(R_t - T_t)}$$

$\boldsymbol{R}_t$ vector of mean % API dissolved for Reference product at time point t

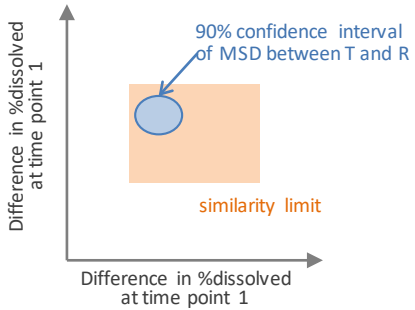
$\boldsymbol{T}_t$ vector of mean % API dissolved for Test product at time point t

$\Sigma_{pooled} = \frac{\Sigma_T + \Sigma_R}{2}$ covariance matrix

- Sum of relative differences, where at each dissolution time point is related to the variability at this time point
- Both the mean profile and the variability are considered

MSD Method

- Determine the similarity limits in terms of MSD based on inter-batch differences in dissolution from Reference (standard approved) batches
- Estimate the MSD between the Test and Reference mean dissolutions
- Estimate 90% confidence interval of true MSD between Test and Reference batches
- Dissolution profiles of the Test and Reference are considered similar if the upper limit of the confidence interval is less than or equal to the similarity limit.



Case Study 1

f2 bootstrapping for highly variable dissolution data

- Drug A is a locally-acting drug
- Product-specific guidance (PSG) recommends in vitro study as one of the options to establish BE, which is comprised of comparative dissolution testing under a range of pH conditions
- At least 12 units each of the Test and Reference products should be tested

pH Condition #1

Product	% Dissolved	Collection Times (min)					
		5	10	20	30	45	60
Test 12 Units	Mean	2	12	53	93	105	106
	Range	1 - 5	5 - 21	22 - 79	69 - 106	98 - 110	99 - 112
	%CV	47.5	42.2	29.4	11.3	4.0	4.2
Reference 12 Units	Mean	3	12	46	95	105	107
	Range	2 - 3	6 - 17	24 - 72	79 - 105	103 - 107	105 - 108
	%CV	20.9	29.2	31.5	8.8	0.9	0.7

High within-batch variability of drug release at early time points (%CV >20%)
→ f2 testing using mean profiles was not applicable

pH Condition #1

Highly variable dissolution data
- f2 bootstrapping was applied for the dissolution profile comparison

	f2 bootstrap mean	Lower 90% CI (5% percentile)
Test vs. Reference (N=12)	66.82	52.87
Reference vs. Reference (N=6)	65.79	47.24

→ dissolution profiles of T and R considered similar

pH Condition #2

Product	% Dissolved	Collection Times (min)					
		5	10	20	30	45	60
Test 12 Units	Mean	8	36	86	103	104	104
	Range	3 to 16	20 to 54	74 to 101	98 to 106	99 to 106	98 to 106
	%CV	57.9	29.1	8.5	1.9	1.7	2.0
Reference 12 Units	Mean	4	22	81	102	104	104
	Range	3 to 8	18 to 41	75 to 93	98 to 104	103 to 104	103 to 106
	%CV	34.6	28.0	6.3	1.6	0.5	0.8

	f2 bootstrap mean	Lower 90% CI (5% percentile)
Test vs. Reference (N=12)	54.49	44.82
Reference vs. Reference (N=6)	78.04	58.99

→ dissolution profiles of T and R are not similar

Case Study 1

- **Conclusion**: For pH condition #2, lower bound of 90% confidence interval for bootstrapping f2 comparing Test vs. Reference is lower than those comparing the Reference against itself under the same condition

  - Dissolution profiles are not comparable between Test and Reference products

- **Option**: Repeat comparative dissolution testing on the proposed test product with a larger sample size to provide a better estimate of the mean difference.

  The dissolution testing should be conducted on at least 24 units (more if necessary) of the Test product and at least two lots of the unexpired Reference product (12 units per lot)

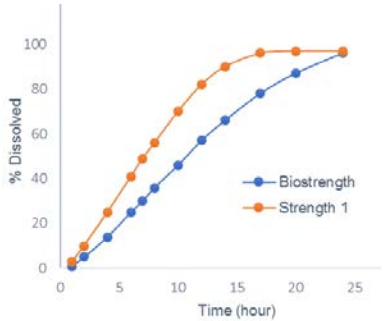
Case Study 2

Applying both f2 bootstrapping and MSD method

- Drug B is an extended-release drug with multiple strengths
- PSG recommends in vivo studies on the middle strength (biostrength), and multimedia dissolution testing under a range of pH conditions as one of the criteria for the waiver request of other lower and higher strengths (non-biostrength)

Strength 1, pH Condition #1

Product	% Dissolved	Collection Times (hour)											
		1	2	4	6	7	8	10	12	14	17	20	24
Biostrength 12 Units	Mean	1	5	14	25	30	36	46	57	66	78	87	96
	%CV	45.23	23.29	17.85	14.08	12.70	12.11	10.04	7.90	6.93	5.24	3.96	1.77
Strength 1 (Non-biostrength) 12 Units	Mean	3	10	25	41	49	56	70	82	90	96	97	97
	%CV	27.83	17.75	14.18	10.04	9.70	8.93	8.45	8.06	6.45	2.06	0.74	0.82



High within-batch variability of drug release at early time points

- Model (Weibull) dependent approach was proposed by applicant but was considered inappropriate

Case Study 2

*"In instances where within batch variation is more than 15% CV, a **multivariate model independent procedure** is more suitable for dissolution profile comparison."*

- Multivariate model independent approach MSD method was then applied
- Bootstrapping f2 and was also applied

Bootstrapping f2

	f2 bootstrap mean	5% percentile
Biostrength vs. Strength 1 (N=12)	37.38	34.58

dissolution profiles of T and R are not similar

MSD method

Upper 90% CI of MSD	Similarity limit (Maximum MSD)
10.65	34.70

dissolution profiles of T and R are similar

- Bootstrapping f2 may be relatively conservative compared to MSD method
- Biowaiver request should be supported by the totality of the submitted information including dissolution profile similarity

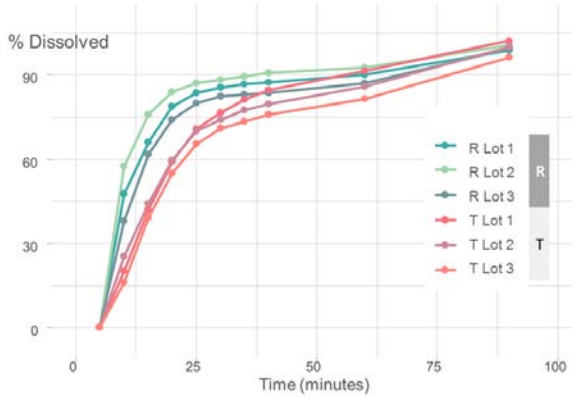
Other deficiencies were identified in this case
→ Option: reformulate the Test product and repeat the dissolution testing

Case Study 3

Comparison of multiple lots of Test and Reference products

- Drug C is an immediate release drug
- PSG recommends in vitro study as one of the options to establish BE, which is to compare **three** lots of Test with **three** lots of Reference products using an optimized QCRT method

Product	% Dissolved	Collection Times (min)									
		5	10	15	20	25	30	35	40	60	90
Test Lot 1 12 Units	Mean	0	20	42	59	71	76	81	85	91	102
	%CV	N/A	39	16	10	8	7	8	8	8	2
Test Lot 2 12 Units	Mean	0	25	44	60	70	74	77	80	86	100
	%CV	N/A	25	16	10	10	9	8	8	7	1
Test Lot 3 12 Units	Mean	0	16	39	55	66	71	74	76	81	96
	%CV	N/A	71	22	12	10	7	8	7	7	2
Reference Lot 1 12 Units	Mean	0	48	66	79	84	86	87	87	90	99
	%CV	N/A	22	10	7	5	5	5	4	4	1
Reference Lot 2 12 Units	Mean	0	57	76	84	87	88	89	91	93	104
	%CV	N/A	16	8	6	5	4	5	5	5	1
Reference Lot 3 12 Units	Mean	0	38	62	74	80	82	83	84	87	99
	%CV	N/A	23	15	11	8	7	6	6	6	2



Case Study 3

I. Pooled data comparison: 36 T vs. 36 R

High Variability

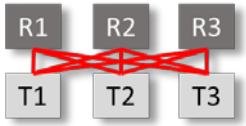
%CV	10 min	15 min
T	25.391	13.597
R	46.134	18.386

Bootstrapping f2

	f2 bootstrap mean	5% percentile
Test vs. Reference (N=36)	36.29	32.87
Reference vs. Reference (N=18)	71.61	56.74

→ dissolution profiles of T and R are not similar

II. Pairwise lot comparison: 12 T vs. 12 R



Comparison (N=12)	f2 bootstrap mean	5% percentile
T1 vs. R1	37.63	32.89
T1 vs. R2	28.92	25.23
T1 vs. R3	**47.38**	**40.45**
T2 vs. R1	39.47	34.77
T2 vs. R2	30.75	27.11
T2 vs. R3	**49.84**	**41.92**
T3 vs. R1	33.63	29.43
T3 vs. R2	26.41	22.94
T3 vs. R3	41.86	35.50
R1 vs. R2	58.46	47.97
R1 vs. R3	65.05	52.02
R2 vs. R3	**46.04**	**38.86**

→ T1 and T2 are comparable to R3

- MSD method also showed lack of similarity
- Results from both methods indicate that the dissolution profiles between Test and Reference are not similar
- Option: Develop new discriminatory QCRT method or reformulate the Test product

Summary

- Guidances reflect the agency's current recommendations for dissolution profile similarity testing
- f2 is a basic tool for dissolution profile similarity assessment
- In case of high variability in the dissolution profiles, appropriate statistical method(s) should be applied to evaluate the dissolution profile similarity
  - f2 bootstrapping
  - MSD method
- Other methods with sufficient justification may be acceptable. Potential applicants are highly encouraged to discuss alternative approaches with OGD

Guidances and References

- Guidance for Industry: Dissolution Testing of Immediate Release Solid Oral Dosage Forms (August 2017)
- Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System. Guidance for Industry (December 2017)
- SUPAC-IR: Immediate-Release Solid Oral Dosage Forms: Scale-Up and Post-Approval Changes: Chemistry, Manufacturing and Controls, In Vitro Dissolution Testing, and In Vivo Bioequivalence Documentation (November 1995)
- SUPAC-MR: Modified Release Solid Oral Dosage Forms Scale-Up and Postapproval Changes: Chemistry, Manufacturing, and Controls; In Vitro Dissolution Testing and In Vivo Bioequivalence Documentation (September 1997)
- Extended Release Oral Dosage Forms: Development, Evaluation, and Application of In Vitro/In Vivo Correlations (September 1997)
- Product Specific Guidances for Generic Drug Development: https://www.fda.gov/drugs/guidances-drugs/product-specific-guidances-generic-drug-development
- Sathe, P. M., Tsong, Y., & Shah, V. P. (1996). In-vitro dissolution profile comparison: statistics and analysis, model dependent approach. Pharmaceutical research, 13(12), 1799-1803.
- Tsong, Y., Sathe, P. M. & Shah, V. P.(2003) In vitro dissolution profile comparison. In Encyclopedia of Biopharmaceutical Statistics, Second Edition, Chow, S.-C., Ed.; CRC Press
- 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. Eur J Pharm Biopharm 112, 67-74.
- M-CERSI Workshop: In Vitro Dissolution Profiles Similarity Assessment in Support of Drug Product Quality: What, How, and When. (May 2019). University of Maryland, Baltimore

Acknowledgments

FDA/OMPT/CDER

OGD/ORS/DQMM

Meng Hu, Ph.D.
Fang Wu, Ph.D.
Youssef Mousa, Ph.D.
Lanyan (Lucy) Fang, Ph.D.
Liang Zhao, Ph.D.

OGD/ORS

Ping Ren, Ph.D.
Lei Zhang, Ph.D.
Robert Lionberger, Ph.D.

OGD/OB

Zhen Zhang, Ph.D.
Leah Falade, Ph.D.
Qing Liu, Ph.D.

Colleagues from OB and OPQ involved in the relevant ANDA reviews

Supported by the Generic Drug User Fee Amendments

Challenge Question #1

Which of the following statements is NOT true?

A. f2 = 100 indicates that the mean difference in the amount dissolved at each time point is 0

B. Dissolution profiles are considered similar if the upper bound (95% percentile) of 90% CI for bootstrapped f2 ≥ 50

C. 3 or more dissolution time points are needed for f2 calculation

D. MSD method is a model independent approach for dissolution profile similarity assessment


FDA
U.S. FOOD & DRUG
ADMINISTRATION