FDA's Current Practice and Challenges in the Use of Dissolution Similarity Testing for Demonstration of Bioequivalence – Case Studies

Zhen Zhang, Ph.D.
Division of Bioequivalence I
Office of Bioequivalence
Office of Generic Drugs, CDER, FDA

Dissolution Similarity Workshop
University of Maryland School of Pharmacy, May 21-22, 2019

Disclaimer

The views and opinions presented here represent those of the speaker and should not be considered to represent advice or guidance on behalf of the U.S. Food and Drug Administration

Outline

- Definition of Dissolution Similarity
- Dissolution Similarity in Bioequivalence Determination
- Office of Bioequivalence's Current Practice and Challenges: Four Case Studies
- Summary/Challenges

Definition of Dissolution Similarity

- Dissolution profiles may be considered similar by virtue of overall profile similarity (e.g., f2 ≥ 50)[1, 2, 3, 4, 5] and similarity at every dissolution sample time point (e.g., ≤ 15%) [1, 3, 4]. – **Profile Comparison**

- When both test and reference products dissolve 85 percent or more of the label amount of the drug in 15 minutes using all three dissolution media, the profile comparison with an f2 test is unnecessary[5]. – **Point Comparison**

- To allow the use of mean data, the coefficient of variation should not be more than 20 percent at the earlier time points (e.g., 15 minutes), and should not be more than 10 percent at other time points[1, 5]. – **Low variability**

Dissolution Similarity in Bioequivalence Determination - Profile Comparison

	Common Medium	Sampling time	Criteria
'waiver' of non-bio strength	QC medium for IR; QC + multimedia for MR	Sufficient number of intervals to characterize the entire dissolution profile of drug product	**Low variable data**: Similar if f2 ≥ 50; **Highly variable data**: Other methods (e.g. bootstrap f2)
Multimedia dissolution for MR products	pH 1.2, 4.5 and 6.8 buffer	Include early sampling times of 1, 2, and 4 hours and continue every 2 hours until at least 80% of the drug is released	
Multimedia dissolution for locally acting drugs	Per PSG	Example: PSG for Mesalamine DR Tablets, 800 mg strength: 0, 10, 20, 30, 45, 60, 75, 90, 120, 150, 180, 240, 300, and 360 minutes or as needed	
QCRT (e.g. Icosapent Ethyl Capsule)	Develop a discriminatory QCRT method	Early times (e.g. 5, 10, 15, 20, 25 minutes) and as frequently as possible, until at least 80% of the drug is released	
DESI Drug	QC medium	Sufficient number of intervals to characterize the entire dissolution profile of drug product	
Half tablets (e.g. scored ER tablets)	Same as whole tablets	Same sampling time points as whole tablets	

QC: Quality Control; **MR**: Modified-Release; **ER**: Extended-Release; **IR**: Immediate Release
PSG: Product-Specific Guidance; **QCRT**: Quantitative Capsule Rupture Test; **DESI**: Drug Efficacy Study Implementation

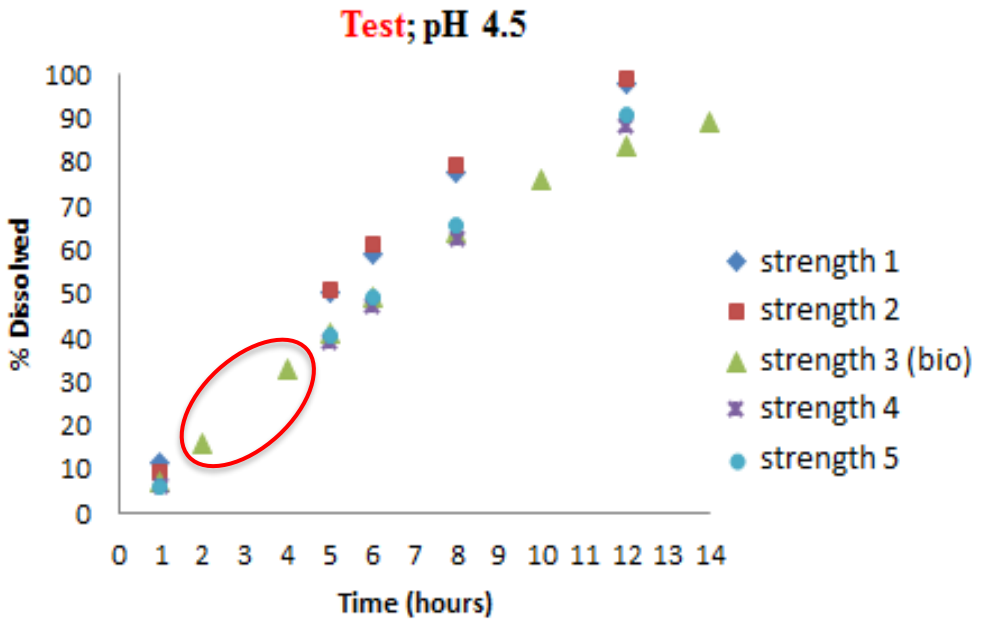
Dissolution Similarity in Bioequivalence Determination - Point Comparison

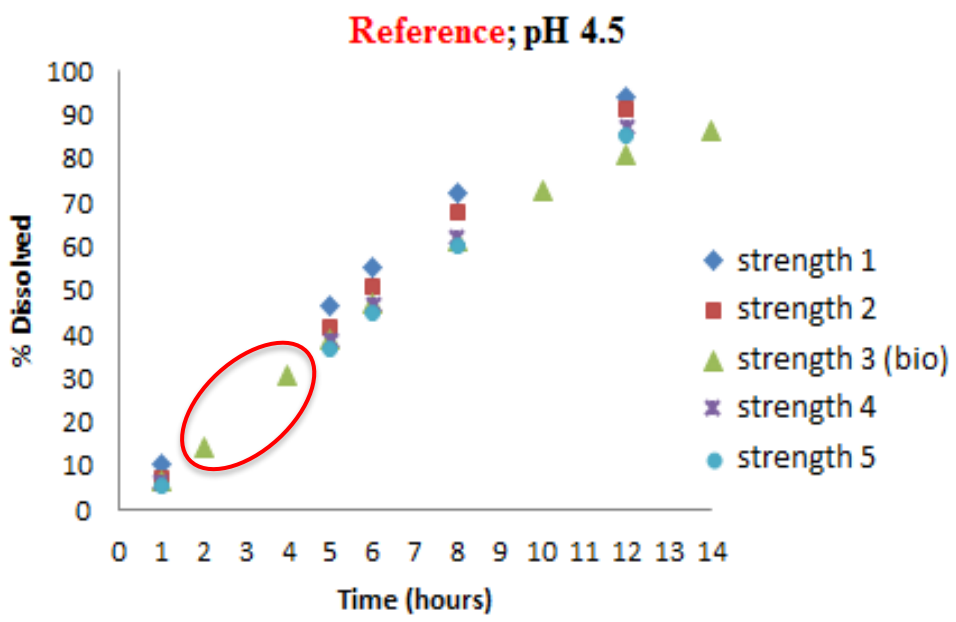
	Common Medium	Sampling Time	Criteria
BCS I/III waiver	1) 0.1 N HCl or SGF w/o enzyme; 2) pH 4.5 buffer; 3) pH 6.8 buffer or SIF w/o enzyme)	e.g., 5, 10, 15, 20, and 30 minutes	**Point Comparison** BCS I: ≥ 85% (mean) within 30 min BCS III: ≥ 85% (mean) within 15 min **Profile Comparison** similar dissolution profiles (e.g. f2 ≥ 50 if applicable)
Alcohol dose dumping	900 mL, 0.1 N HCl, USP apparatus 2 at 50 rpm, w/ or w/o Alcohol	Samples of the media are taken once every 15 minutes until 2 hours is reached	Comparable % dissolved drug product

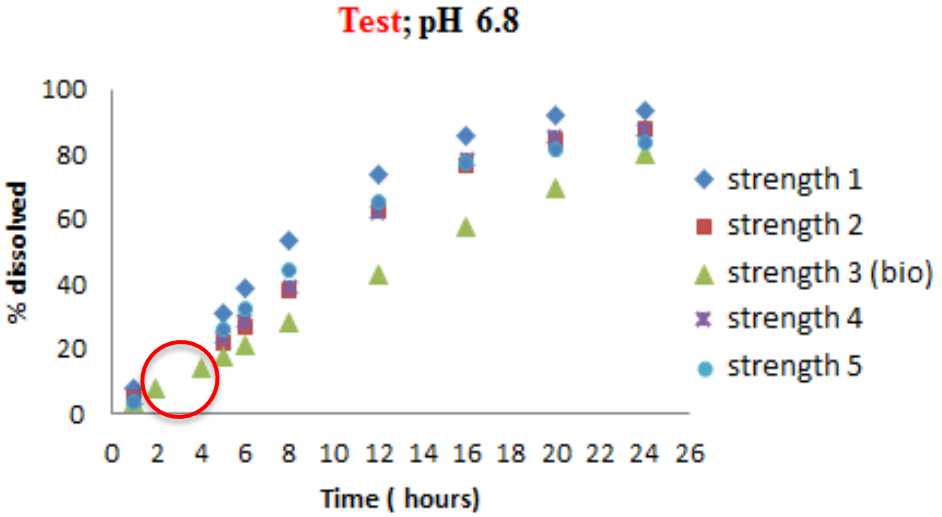
BCS: Biopharmaceutics Classification System; **SGF**: Simulated Gastric Fluid; **SIF**: Simulated Intestinal Fluid;
QC: Quality Control; **IR**: Immediate Release

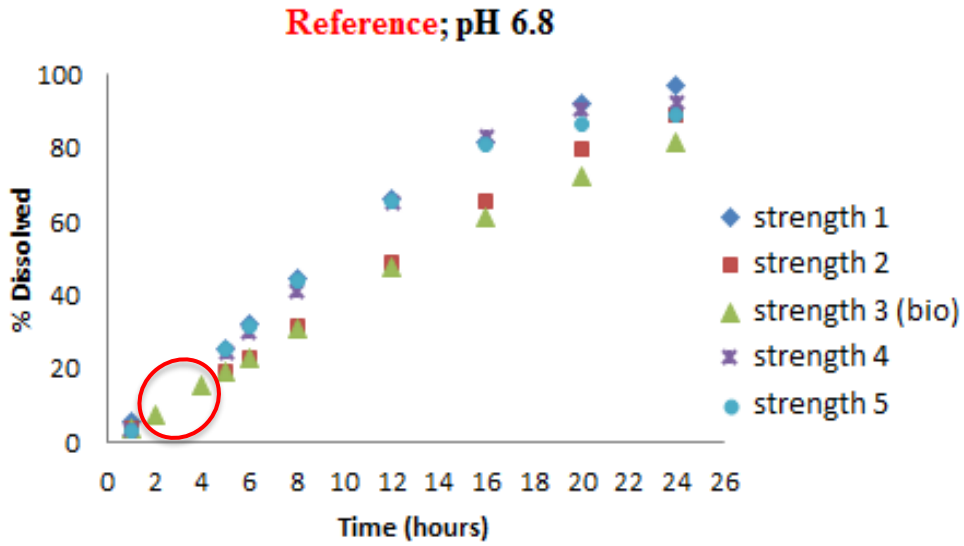
Case Study #1: Justification of Missing Sampling Time

- The drug product A is an extended release tablet, which has five strengths
- The PSG recommends in vivo bioequivalence (BE) studies on the middle strength.
- As one of the criteria to evaluate the waiver request of non-bio strengths, the PSG recommends multimedia dissolution testing at pH 1.2, 4.5 and 6.8 buffers including early sampling times of 1, 2 and 4 hours and continue every 2 hours until at least 80% of the drug is release.









Justification: 1) low variability; 2) even though non-bio strengths do not have release data at 2-hour and 4-hour time points, the dissolution data include early, middle, and complete release at different time points, which is sufficient to capture the whole release profile.

Challenge: What is the appropriate sampling time for immediate release solid oral dosage forms, especially when the drug release does not reach 85% within 15 min? Include early times (e.g., 5, 10, 15, 20, 25 minutes) and as frequently as possible?

Case Study #2: Justification of Similarity When f2 < 50

- Same drug product and dissolution data as in Case Study #1;
- Acceptable in vivo BE studies on the middle strength (bio-strength) and formulation proportionality across all strengths;
- Per Guidance for Industry: Bioequivalence Studies with Pharmacokinetic Endpoints for Drugs Submitted Under an ANDA (PK BE Guidance; Dec 2013), *we recommend that the drug products exhibit similar dissolution profiles* ***between the strength on which BE testing was conducted and other strengths based on the f2 test in at least three dissolution media (e.g., pH 1.2, 4.5, and 6.8).***

F2 metric, biostudy strengths compared to other strengths of test product

Biostudy Strength	Other Strength	QC media	pH 1.2	pH 4.5	pH 6.8
Bio-strength	Strength 1	61.15	71.20	49.45	33.88
Bio-strength	Strength 2	63.96	82.83	47.13	45.96
Bio-strength	Strength 4	70.76	56.06	71.45	44.61
Bio-strength	Strength 5	71.96	52.93	66.65	43.8

F2 Metric Test vs Reference (all strengths)

Strength(s)	QC media	pH 1.2	pH 4.5	pH 6.8
Strength 1	56.69	57.09	69.36	62.59
Strength 2	59.22	61.16	52.72	57.68
Bio-strength	46.58	69.48	80.33	76.49
Strength 4	79.91	73.01	97.1	73.2
Strength 5	72.15	70.11	68.24	76.5

F2 metric, biostudy strengths compared to other strengths of reference product

Biostudy Strength	Other Strength	QC media	pH 1.2	pH 4.5	pH 6.8
Bio-strength	Strength 1	49.36	65.52	51.01	41.53
Bio-strength	Strength 2	53.46	83.71	60.56	68.79
Bio-strength	Strength 4	58.28	58.55	68.2	43.83
Bio-strength	Strength 5	50.57	53.79	69.62	45.15

Conclusion: The dissolution data are acceptable to support waiver request.
Note: May not be applicable to other cases.

Challenges: F2 ≥ 50 is used as a criterion to determine the dissolution similarity.

1) If f2 fails within a close margin (e.g., f2 = 49), what is the likelihood of rejecting a 'good' drug product if we strictly follow the criterion of f2 ≥ 50?
2) If f2 fails largely (e.g., f2 = 40), what is an acceptable justification, e.g., clinical relevance, comparable dissolution profile between test and reference?

Case Study #3: Bootstrap f2 for Highly Variable Dissolution Data

- Drug product B is a locally acting, extended-release drug;
- The PSG recommends in vitro comparative dissolution testing at different conditions as one of the pivotal BE studies.

Test Condition #1:

		Collection Time Points						
		1 hrs.	2 hrs.	4 hrs.	6 hrs.	8 hrs.	10 hrs.	12 hrs.
Test 12 units/lot	Mean	3%	21%	47%	67%	85%	95%	97%
	Range	2-5%	19-22%	43-52%	56-77%	73-93%	90-97%	95-98%
	% RSD	23.21%	4.76%	6.91%	10.63%	6.87%	2.05%	0.82%
Reference 12 units/lot	Mean	1%	13%	45%	84%	94%	95%	95%
	Range	1-3%	6-22%	25-70%	53-94%	91-96%	94-96%	94-97%
	% RSD	47.19%	32.70%	26.34%	16.70%	1.29%	0.63%	0.83%

To allow use of mean data, the percent coefficient of variation at the earlier time points (e.g., 15 minutes) should not be more than 20%, and at other time points should not be more than 10%.

Bootstrap f2

	F2_boot_mean	F2_5th
Test vs Reference	51.29	45.58
Reference vs Reference	57.32	41.08

→ Similar T vs. R

Test Condition #2:

		Collection Time Points						
		1 hrs.	2 hrs.	4 hrs.	6 hrs.	8 hrs.	10 hrs.	12 hrs.
Test 12 units/lot	Mean	0%	1%	27%	47%	63%	79%	92%
	Range	0%	0-3%	24-36%	41-64%	52-85%	65-96%	81-98%
	% RSD	NA	86.50%	11.83%	12.84%	13.93%	11.61%	5.12%
Reference 12 units/lot	Mean	0%	1%	7%	28%	63%	88%	91%
	Range	0%	0-2%	0-15%	0-47%	1-96%	1-100%	2-101%
	% RSD	NA	69.28%	64.57%	45.46%	40.28%	31.61%	30.83%

Bootstrap f2

	F2_boot_mean	F2_5th
Test vs Reference	44.85	41.12
Reference vs Reference	52.27	29.99

→ Different T vs. R

Conclusion: For Test Condition #2, dissolution profiles are not comparable between the test and reference products. Repeat comparative dissolution testing on the unexpired test product using 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 unexpired test product and at least two lots of unexpired reference product (12 units per lot)

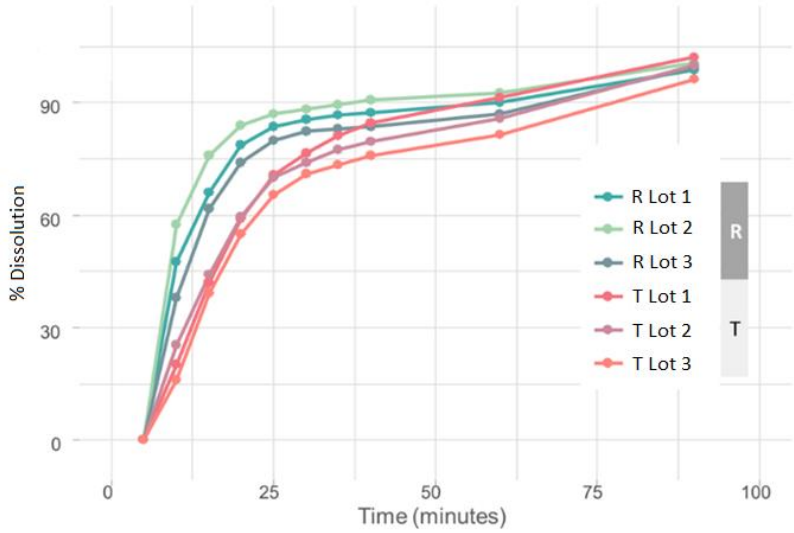
Challenges:

1) When are the dissolution data considered as highly variable for modified release drug product? Do we need to take the drug release range at early time points into consideration?
2) Are 12-unit data sufficient for the comparison of dissolution profiles for highly variable dissolution data?

Case Study #4: Comparison of 3 Test Lots and 3 Reference Lots

- Drug C is an immediate release drug product
- The PSG recommends to compare three lots of the test product with three lots of the reference product using an optimized QCRT method

		Collection Times									
		5 min	10 min	15 min	20 min	25 min	30 min	35 min	40 min	60 min	90 min
Test Lot 1 12 units	Mean	0	20	42	59	71	76	81	85	91	102
	Range	0-0	3-35	34-56	51-70	64-79	69-84	72-90	74-93	79-101	98-105
	%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
	Range	0-0	16-39	31-61	50-73	63-89	68-91	71-92	72-93	75-97	98-102
	%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
	Range	0-0	1-36	23-54	43-65	54-74	61-79	65-82	68-84	73-87	92-98
	%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
	Range	0-0	27-68	58-80	71-86	76-89	78-91	79-91	80-92	83-94	97-100
	%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	100
	Range	0-0	42-70	66	8577-93	81-95	81-95	81-96	82-97	83-98	99-102
	%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
	Range	0-0	22-55	47-79	60-88	69-90	73-91	74-91	74-92	77-95	97-103
	%CV	N/A	23	15	11	8	7	6	6	6	2



Individual Test Lot vs. Individual Reference Lot (12 T vs. 12 R) – Bootstrap f2

Comparison	f2_boot_mean	f2_5th
R1 vs. R2	58.45909	47.97313
R1 vs. R3	65.04528	52.01843
R2 vs. R3	46.04056	38.86425
T1 vs. R1	37.62635	32.89369
T1 vs. R2	28.91628	25.23499
T1 vs. R3	**47.37881**	**40.45291**
T2 vs. R1	39.46997	34.77204
T2 vs. R2	30.75399	27.10914
T2 vs. R3	**49.84414**	**41.9163**
T3 vs. R1	33.62667	29.43244
T3 vs. R2	26.4102	22.9441
T3 vs. R3	41.85864	35.50467

→ **Results**: Only Test Lot 1 and Test Lot 2 are comparable to Reference Lot 3.

Pooled Test Data vs. Pooled Reference Data (36 T vs. 36 R) – Bootstrap f2

Comparison	f2_boot_mean	f2_5th
Test vs. Reference	36.29	32.87
Reference vs. Reference	71.61	56.74

→ **Results**: The QCRT data are not comparable between the test and reference products

Exploratory: Mahalanobis distance (M-distance)-based approach also shows that the QCRT data are not comparable between the test and reference products

Challenge: What is an appropriate approach to compare multiple T vs. multiple R? Individual or pooled data?

Summary

- Dissolution is a critical tool for the evaluation of generic drug products;
- Low variable dissolution data
  - Similar if f2 ≥ 50;
- Highly variable dissolution data
  - Bootstrap f2 method;
  - Other methods with sufficient justification are also acceptable.
- If f2 fails to meet the acceptance criteria, justification is welcome.
- Challenges

Challenges

- What is the appropriate sampling interval for immediate release solid oral dosage forms, especially when the drug release does not reach 85% within 15 min? Include early times (e.g., 5, 10, 15, 20, 25 minutes) and as frequently as possible?

- If f2 fails within a close margin (e.g., f2 = 49), what is the likelihood of rejecting a 'good' drug product if we strictly follow the criterion of f2 ≥ 50?

- If f2 fails largely (e.g., f2 = 40), what is an acceptable justification, e.g., clinical relevance, comparable dissolution profile between test and reference?

- When are the dissolution data considered as highly variable for modified release drug products? Do we need to take the drug release range at early time points into consideration?

- Are 12-unit data sufficient for the comparison of dissolution profiles for highly variable dissolution data?

- What is an appropriate approach to compare multiple T vs. multiple R? Individual or pooled data?

Acknowledgements

Division of Bioequivalence I
Bing V. Li
Utpal Munshi
Qing Liu
Rong Wang

Division of Bioequivalence II
Hongling Zhang

Division of Bioequivalence III
Nilufer Tampal
April Braddy

Division of Biopharmaceutics
Sandra Suarez

References

1. **Dissolution Testing of Immediate Release Solid Oral Dosage Forms (Aug 1997)**: Dissolution profiles may be considered similar by virtue of (1) overall profile similarity and (2) similarity at every dissolution sample time point… Generally, f1 values up to 15 (0-15) and f2 values greater than 50 (50-100) ensure sameness or equivalence of the two curves… To allow use of mean data, the percent coefficient of variation at the earlier time points (e.g., 15 minutes) should not be more than 20%, and at other time points should not be more than 10%.
2. **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 (Nov 1995)**: An f2 value between 50 and 100 suggests the two dissolution profiles are similar.
3. **Extended Release Oral Dosage Forms: Development, Evaluation, and Application of In Vitro/In Vivo Correlations (Sep 1997)**: A model independent approach using a similarity factor, and comparison criteria are described in SUPAC-MR.
4. **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 (Sep 1997)**: An f2 value between 50 and 100 suggests the two dissolution profiles are similar. Also, the average difference at any dissolution sampling time point should not be greater than 15% between the changed drug product and the biobatch or marketed batch (unchanged drug product) dissolution profiles. An f2 value less than 50 does not necessarily indicate lack of similarity.
5. **Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System. Guidance for Industry (Dec 2017)**: Two dissolution profiles are considered similar when the f2 value is ≥ 50. To allow the use of mean data, the coefficient of variation should not be more than 20 percent at the earlier time points (e.g., 15 minutes), and should not be more than 10 percent at other time points. Only one measurement should be considered after 85 percent dissolution of both products. In addition, when both test and reference products dissolve 85 percent or more of the label amount of the drug in 15 minutes using all three dissolution media recommended above, the profile comparison with an f2 test is unnecessary.


FDA
U.S. FOOD & DRUG
ADMINISTRATION