

# EMA (2018) Q&A on Mahalanobis distance (MD) to assess drug dissolution profiles – Statistical critique & demonstration of adequacy of T2EQ approach (based on MD)

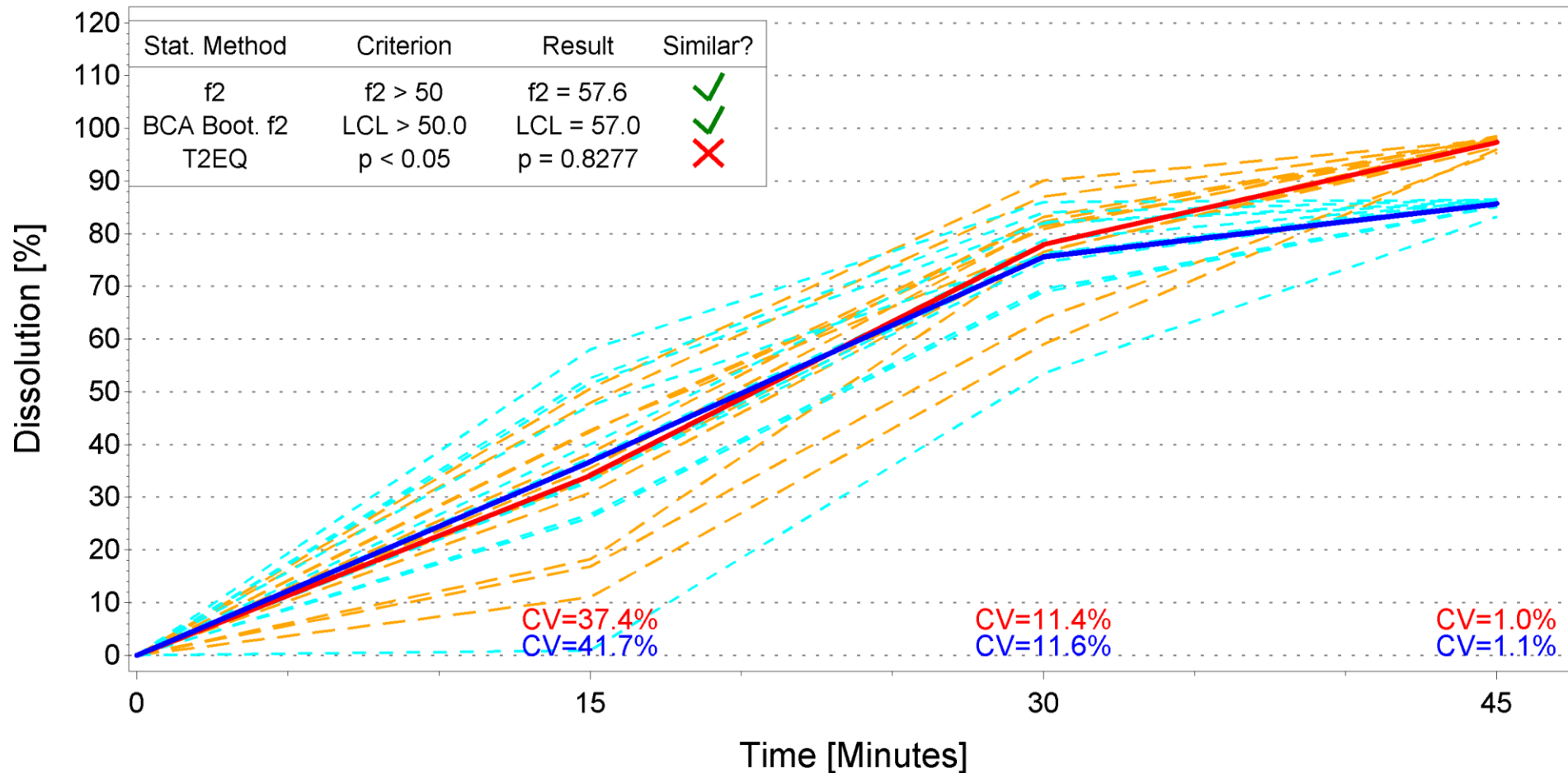
Dr. Thomas Hoffelder

4th EFSPi Workshop on Regulatory Statistics  
September 23-24, 2019

# Dissolution profile equivalence - example (source of figure: Hoffelder, 2018)

Figure 2

Reference versus Test\_2  
Simulation Example



Symbols:   
-- Individual Profiles of Reference — Mean Profile (n=12) of Reference   
-- Individual Profiles of Test\_2 — Mean Profile (n=12) of Test\_2

# Topic 1: “Average difference <10%” or “maximum difference <10%”?

## EMA (2010) guideline:

- *“The similarity acceptance limits should be pre-defined and justified and not be **greater than a 10% difference**. ... An  $f_2$  value between 50 and 100 suggests that the two dissolution profiles are similar.”*

Note:  $f_2$ -method refers to the **average difference**. I.e., EMA (2010) - like FDA (1997) - refer to the **average difference of disso profiles**.

## EMA (2018) Q&A document:

1. **contains** two statements, which refer to the **maximum difference** across all timepoints as the similarity criterion

➔ may be interpreted as a **new (!), much stricter** regulatory acceptance criterion for similarity of disso profiles!

Note: Unusual for a Q&A document to revise a major aspect of the guideline.

2. **recommends bootstrapped  $f_2$  as “preferred method”**

➔ bootstrapped  $f_2$  targets the average (NOT maximum!) difference criterion!

➔ confusion for regulatory assessors and the industry

Remark: Comparison **bootstrapped  $f_2$**  vs. **T2EQ** approach (see Hoffelder, 2019a)

- T2EQ approach seems to outperform bootstrapped  $f_2$ .
- bootstrapped  $f_2$ : poor type 1 error control

## Topic 2: Although T2EQ is based on MD, EMA (2018)'s concerns do NOT apply here

**Mahalanobis distance:**

$$MD = (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2)$$

**MD point estimate:**

$$\widehat{MD} = (\hat{\mu}_1 - \hat{\mu}_2)' S^{-1} (\hat{\mu}_1 - \hat{\mu}_2)$$

**S:** empirical covariance matrix of pooled sample

**Equivalence hypotheses:**

$$H_0 : MD \geq EM \quad \text{versus} \quad H_1 : MD < EM$$

**EM:** Equivalence margin (acceptance criterion)

EMA (2018) generally discourages the MD based methods.

- seems to be based on the misconception, that MD is always connected to a fixed EM
  - Concerns against MD with fixed EMs: MD shrinks when the variance of the experiment increases
    - the higher the variability the higher is the power
    - poorly designed experiments would be rewarded
    - not compliant with EMA (2010) because a decision in favor of equivalence might be possible in spite of a profile difference > 10% at all time points
  - **T2EQ approach (not based on a fixed EM !)**
    - + T<sup>2</sup>-test for equivalence according to Wellek (2010)
    - + EM according to Tsong et al. (1996)
    - + Restrictions according to EMA (2010)
- $EM = D' S^{-1} D$ , where  $D = (10, 10, \dots, 10)$   
→ analogy to f2

## Topic 2: Although T2EQ is based on MD, EMA (2018)'s concerns do NOT apply here

**Mahalanobis distance:**

$$MD = (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2)$$

**MD point estimate:**

$$\widehat{MD} = (\hat{\mu}_1 - \hat{\mu}_2)' S^{-1} (\hat{\mu}_1 - \hat{\mu}_2)$$

**S:** empirical covariance matrix of pooled sample

**Equivalence hypotheses:**

$$H_0 : MD \geq EM \quad \text{versus} \quad H_1 : MD < EM$$

**EM:** Equivalence margin (acceptance criterion)

EMA (2018) generally discourages the MD based methods.

- seems to be based on the misconception, that MD is always connected to a fixed EM
  - Concerns against MD with fixed EMs: MD shrinks when the variance of the experiment increases
    - the higher the variability the higher is the power
    - poorly designed experiments would be rewarded
    - not compliant with EMA (2010) because a decision in favor of equivalence might be possible in spite of a profile difference > 10% at all time points
  - **T2EQ approach (not based on a fixed EM !)**
    - + T<sup>2</sup>-test for equivalence according to Wellek (2010)
    - + EM according to Tsong et al. (1996)
    - + Restrictions according to EMA (2010)
- }  $EM = D' S^{-1} D$ , where  $D = (10, 10, \dots, 10)$   
→ analogy to f2

## Topic 2: Although T2EQ is based on MD, EMA (2018)'s concerns do NOT apply here

**T2EQ approach not affected by the concerns against MD because**

- the covariance matrix enters into both sides of the equivalence hypotheses, into **MD as well as into EM.**

Increased variance → **EM decreases in the same way as MD.**

- **EM defined by a shift in location of 10%** at all dissolution time points.  
→ **T2EQ compliant with the EMA guideline** independent of the present covariance matrix
- Remark: Simulations in Hoffelder (2019a): increased variability reduces the power

- + T2EQ: statistically sound, well documented and **fully compliant with EMA (2010).**
- + T2EQ also supports an interpretation in **analogy to f2.**
- + **Recommendation:** put the EMA (2018) statements about MD in the appropriate context and also to extend the EMA (2018) statements to cover other innovative and scientifically sound inferential procedures, such as T2EQ.

# Conclusions/Questions

- 1) We recommend to put the EMA (2018) statements about MD in the appropriate context and also to extend the EMA (2018) statements to cover other innovative and scientifically sound inferential procedures, such as T2EQ.**
  
- 2) Inconsistencies regarding acceptance criterion for dissolution profiles within EMA (2018) and between EMA (2018) and EMA (2010)  
Which criterion should be used?  
„average diff. < 10%“      versus      „max. diff < 10%“**

End

---

**Back up slides Hoffelder**



# Literature

- EMA (2010).** *Guideline on the Investigation of Bioequivalence.* London: European Medicines Agency, CHMP.
- EMA (2018).** *Question and answer on the adequacy of the Mahalanobis distance to assess the comparability of drug dissolution profiles.* London: European Medicines Agency.
- FDA (1997).** *Guidance for Industry: Dissolution Testing of Immediate Release Solid Oral Dosage Forms.* Rockville: Food and Drug Administration FDA, CDER.
- Hoffelder, T. (2018).** Comparison of dissolution profiles – a statistician’s perspective. *Therapeutic Innovation & Regulatory Science* **52:4**, 423–429.
- Hoffelder, T. (2019a).** Equivalence analyses of dissolution profiles with the Mahalanobis distance. *Biometrical Journal* **61**, 1120–1137.
- Hoffelder, T. (2019b).** Author response to the Letter to the Editor “Equivalence analyses of dissolution profiles with the Mahalanobis distance: A regulatory perspective and a comparison with a parametric maximum deviation-based approach”. *Biometrical Journal* **61**, 1138–1140.
- Tsong, Y., Hammerstrom, T., Sathe, P., Shah, V. P. (1996).** Statistical assessment of mean differences between two dissolution data sets. *Drug Information Journal* **30**, 1105–1112.
- Wellek, S. (2010).** *Testing Statistical Hypotheses of Equivalence and Noninferiority. Second Edition.* Boca Raton: Chapman & Hall / CRC Press LLC.

# Background of dissolution profile comparisons

- Situation: postapproval change:
  - scale-up changes
  - manufacturing site changes
  - component and composition changes
  - equipment and process changes
- Manufacturer must prove that quality, safety and efficacy are not affected by the change. The question  
“Is the drug product made after the change equivalent to the drug product made before the change?”  
has to be answered.
  - TEST group
  - REF group
- **Dissolution profile comparison** successful → no bioequivalence study necessary

# Details: The similarity factor $f_2$

Standard approach: similarity factor  $f_2$

- $n$ : number of dissolution time points,  $R = (R_1, \dots, R_n)$ : reference mean profile,  $T = (T_1, \dots, T_n)$ : test mean profile

$f_2$  is a transformation of the quadratic mean (over time) of the differences between reference and test mean:

$$QMD = \sqrt{(1/n) \sum_{t=1}^n (R_t - T_t)^2} \Rightarrow f_2 = 50 \log_{10} \left( \frac{100}{\sqrt{1 + QMD^2}} \right)$$

- acceptance criterion:  $f_2 > 50 \Leftrightarrow QMD < \sqrt{99} = 9.95 \approx 10$

→ profiles similar if average difference between profile means is below 10%

- $f_2$  depends on means only, **point estimate, no control of type I error**

→ the less reliable the higher the variability of the underlying data

→ **guideline restrictions for  $f_2$  if variability exceeds certain thresholds**

# Details: The similarity factor $f_2$

## Example:

3 time points:	{	Mean difference at time point nr. 1:	<b>11%</b>
		Mean difference at time point nr. 2:	<b>11%</b>
		Mean difference at time point nr. 3:	<b>3%</b>

With the notation on slide 6, it is

$$QMD = \sqrt{\frac{1}{3} (11^2 + 11^2 + 3^2)} = 9.15 \quad \Rightarrow \quad f_2 = 51.81$$

The **quadratic mean of the differences** is lower than 10 resulting in  $f_2 > 50$

➔ Dissolution profile similarity is assessed although the mean difference between the profiles is above 10% in two of three time points.

# Details: T2EQ approach

- **Distance measure:** Mahalanobis distance (MD)

$$MD = (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2)$$

$\mu_1$  and  $\mu_2$  are the expected values of both reference and test group and  $\Sigma$  is the common covariance matrix

- **Equivalence hypotheses:**

$$H_0 : MD \geq EM \quad \text{versus} \quad H_1 : MD < EM$$

**EM: Equivalence margin (similarity limit / acceptance criterion)**

- **Note:** Sometimes, terminology „Mahalanobis distance“ used for  $\sqrt{MD}$

# Details: T2EQ approach

## Specific choice of the T2EQ EM

### Tsong et al. (1996):

- “**global similarity limit**” for MD in the context of comparing dissolution profiles  
→ product independent, always practically feasible EM.
- EM is defined by a max. allowed constant shift between REF and TEST  $D = (d, d, \dots, d)$
- this max. allowed shift is related to the present variability  $EM = D'S^{-1}D$   
 $S$ : pooled empirical covariance matrix

### EMA (2010):

- “similarity acceptance limits should ... not be greater than a 10% difference”

→  $EM = D'S^{-1}D$ , where  $D = (10, 10, \dots, 10)$

→ EM defined by a difference of 10% at all time points  
**same interpretation as for  $f_2$ !**

**Acceptance criterion: “Average” profile difference < 10%**

# Details: T2EQ approach

## T2EQ approach

- T<sup>2</sup>-test for equivalence according to Wellek (2010)
  - EM according to Tsong et al. (1996)
  - EM restrictions according to EMA (2010)
- } T2EQ approach
- **p-value of the test** can be calculated
    - => no simulations or numerical methods necessary
    - => “p-value < 0.05” is not more difficult to understand as “ $f_2 > 50$ ”
  - test result can be reported via p-value or via upper confidence limit for the Mahalanobis distance

# Analogy between T2EQ and $f_2$

“Average” profile difference < 10%

$f_2$ : **quadratic mean** (over time) of the differences between reference and test mean

**T2EQ: MD based weighted mean** (over time) of the differences between reference and test mean

**Correlation adapted mean** (over time) of the differences between reference and test mean