

Review of the EMEA Guidelines' In-Vitro Equivalence Criteria for Cascade Impaction Data

Dennis Sandell
S5 Consulting

2010-Sep-09 / dennis@s5consulting.com

Outline

- Background
- EU Guidance CPM/ERP/4151/00 Rev. 1
- The Question
- Approach
- Simulation
- One approach for realistic acceptance criteria?
- Conclusions

Background (1)

- How to show equivalence / similarity between two inhalation products?
- Mainly of interest for generics & "device/product changes", but also for development of new products
- Engaged debate...
 - When is clinical trials required?
 - What trials?
 - When is in-vitro data sufficient?
 - What in-vitro tests?
 - What is equivalent? (Acceptance criteria)
- Limited understanding of IVIVC ensures that debate will continue

3

Background (2)

- Regulators are trying to provide some directions
 - FDA 1999/2003: Bioavailability and Bioequivalence Studies for Nasal Aerosols and Nasal Sprays for Local Action
 - FDA 2001: Statistical Approaches to Establishing Bioequivalence
 - CA 2007: Submission Requirements for Subsequent Market Entry Inhaled Corticosteroid Products for Use in the Treatment of Asthma
 - EU 2009: CPMP/EWP/4151/00 Rev. 1 Guideline on the requirements for clinical documentation for orally inhaled products (OIP) including the requirements for demonstration of therapeutic equivalence between two inhaled products for use in the treatment of asthma and chronic obstructive pulmonary disease (COPD) in adults and for use in the treatment of asthma in children and adolescents
 - List is not complete...

4

EU Guidance (1)

- The focus of this presentation is the guidance on what would constitute equivalence for impactor data
- "The maximum allowable *in vitro* difference should be indicated and justified, e.g. $\pm 15\%$ may be justifiable. Per impactor stage or justified group of stages the 90% confidence intervals for the observed *in vitro* differences must be calculated. Based on the pre-established protocol and maximum allowable differences, a decision regarding equivalence can be made."
- What does this mean?

5

EU Guidance (2)

- Overall average for TEST within 85-115% of overall average for REF?
- All batch averages for TEST within 85-115% of overall average for REF?
- All individual unit results within 85-115% of overall average for REF?
- All batch averages for TEST between 85% of lowest and 115% of highest batch average for REF?
- All individual unit results for TEST between 85% of lowest and 115% of highest individual unit result for REF?
- More options exists
- But where does the confidence interval fit in?

6

EU Guidance (3)

- My interpretation:
 - Collect impactor data for TEST and REF products
 - For each stage (or selected groupings) construct 90% confidence interval $I = (LL, UL)$ for μ_{TEST}/μ_{REF} (ratio between the mean amounts for the stage)
 - "Stage equivalence" can be claimed if the confidence interval is contained in the acceptance interval (0.85, 1.15); that is, if $LL \geq 0.85$ and $UL \leq 1.15$
 - "Impactor equivalence" can be claimed if all stages comply
- This is a statistically sound principle to show equivalence
- Guidance opens for other choice of acceptance interval

7

The Question

- Is this a reasonable approach?
- The likelihood to be able to claim equivalence depends on several things
 - How close TEST is to REF (ie true value of μ_{TEST}/μ_{REF})
 - Variability of TEST and REF data
 - Amount data collected
 - Number of stages / groupings
 - Confidence level (given)
 - Maximum allowable difference
- These questions can be studied by simulation

8

Approach

- Compare the reference product to itself using simulation
- Study reasonable range of product variability and test plans
- If a "typical" product fails equivalence to itself something must be changed

9

Simulation - model

- REF: N batches, n results per batch
 $r_{ij} = A_i + e_{ij}$, $i=1, \dots, N$; $j = 1, \dots, n$
 A_i normal with mean μ_{REF} and std σ_{REF} (batch)
 e_i normal with mean 0 and std ξ_{REF} (unit)
- TEST: corresponding
- For selected choice of N, n, means and std's randomly generate data
- Calculate CI and record compliance with acceptance criteria
- Repeat many times

10

Simulation – Fieller interval

- Let r denote the average of the REF results and R denote an estimate for the variance of r .
- Similarly, let t denote the average of the TEST results and T denote an estimate for the variance of t .
- The limits of a 90% Fieller interval for $\mu_{\text{TEST}}/\mu_{\text{REF}}$ is then given by

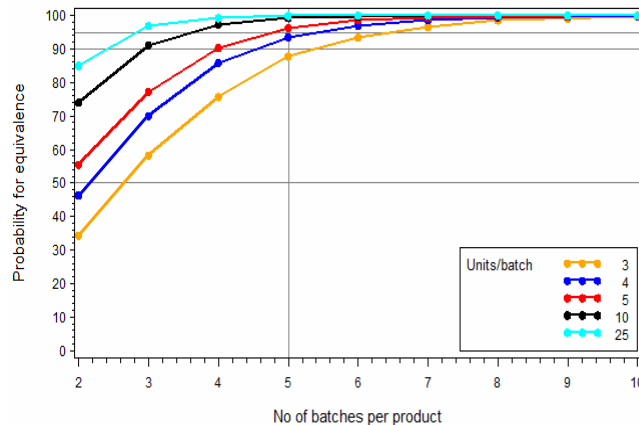
$$\frac{t/r}{k-1} \left[1 \pm 1.6449 \sqrt{(1-k)T + (t/r)^2 R} \right]$$

where $k = (1.6449/r)^2 R$

11

Simulation – Varying sample size

- One stage: $\mu_{\text{TEST}}/\mu_{\text{REF}} = 1$; Batch/unit RSD = 4/10% (TEST = REF)

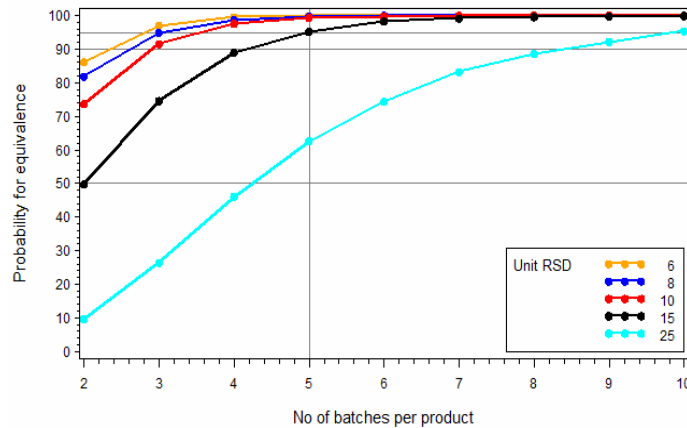


- 10 units for each of 3 batches \Rightarrow 91% pass; 3 units for each of 10 batches \Rightarrow 100% pass
- Clear that increasing number of batches is better than increasing testing per batch
- Many units per batch only worthwhile with few batches

12

Simulation – Varying unit RSD

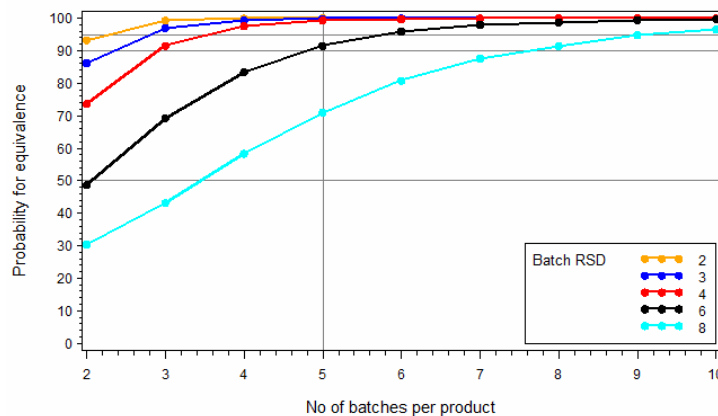
- One stage: $\mu_{\text{TEST}}/\mu_{\text{REF}} = 1$; Batch RSD = 4%; 10 units/batch (TEST = REF)



- Chance to pass decrease with increasing between unit RSD
- With unit RSD < 15% and 10 units/batch 5 batches are enough
- For stages /groupings with low deposition (thus high RSD) many batches are needed for

Simulation – Varying batch RSD

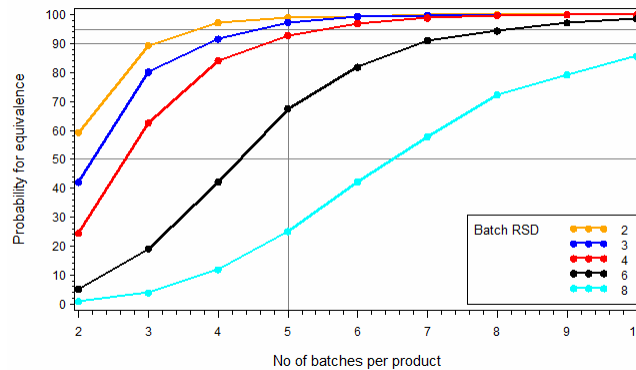
- One stage: $\mu_{\text{TEST}}/\mu_{\text{REF}} = 1$; Unit RSD = 10%; 10 units/batch (TEST = REF)



- Chance to pass decrease with increasing between batch RSD
- With batch RSD > 6% at least 7 batches are needed

Simulation – 4 groupings

- For each added grouping the risk to fail at least one increases
- Assume 10 units per batch and $\mu_{\text{TEST}}/\mu_{\text{REF}} = 1$ for all 4 groupings
- Assume unit RSDs of 6, 8, 10 & 15%; same batch RSD for all groupings



- Unless variabilities are very small and/or testing extensive, it is unlikely that one can show that a product is equivalent to itself

15

Conclusions this far

- For a product with "normal" variabilities and using a "standard" test plan, it is very difficult to show that it is "equivalent to itself"
- It is therefore unlikely that a TEST product would manage this
- There is nothing wrong with the principle for comparing TEST to REF
- What is required is to adjust the acceptance limits

16

One approach for realistic acceptance criteria?

- Select a test plan (i.e., 3 batches, 10 units per batch)
- Select stages / groupings for comparison
- Determine batch and unit RSD for REF
- For each stage / grouping, calculate acceptance limits corresponding to 99% chance to pass
- These are the limits TEST needs to comply with in order to claim "impactor equivalence" to REF (assuming same test plan)

17

Example (1)

- 3 batches, 10 units per batch
- 4 stages / groupings: A, B, C and D
- RSDs:

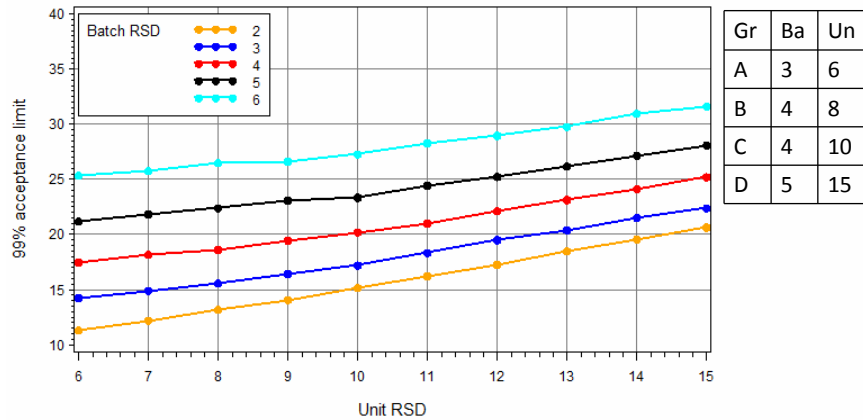
Grouping	Batch	Unit
A	3	6
B	4	8
C	4	10
D	5	15

- Intended to illustrate a realistic situation

18

Example (2)

- 99% acceptance limits for different combinations of batch and unit RSD's (3 batches /10 units test plan)

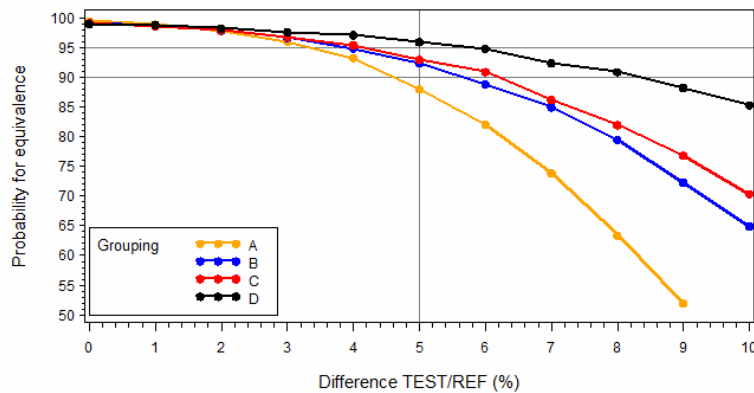


Obtained acceptance limits: A = 14; B = 19; C= 20; D = 27

19

Example (3)

- Probability to show TEST equivalent as function of difference to REF
- Assumes REF and TEST have same variabilities
- Different groupings shown separately



- Reasonable chance to pass for up to 3% difference

20

Conclusions

- EU guidance recommended approach to show "impactor equivalence" has a principally sound interpretation
- 85-115% acceptance limit is too tight as it would be difficult to show that a normal product is equivalent to itself
- It is suggested that "baseline" acceptance limits are calculated based on the selected test plan and performance of the reference product; one procedure for this has been outlined
- Such limits would provide a more reasonable goal for a test product
- Choice in example of 99% as "tuning parameter" need to be discussed
- Test plan should be selected based on power considerations (not shown)
- Same approach can be used for any in-vitro test data comparison

- The questions whether "impactor equivalence" in particular and "in-vitro equivalence" in general can replace "in-vivo equivalence" remains