Comparability study to support commercial process change via stability study

Bianca Teodorescu – EBE, UCB
Cyrille Chéry – EBE, UCB

EMA Workshop " Draft Reflection Paper on statistical methodology for the comparative assessment of quality attributes in drug development"

3-4 May 2018

This is a joint industry presentation on behalf of the trade associations shown

AGENDA

- ✓ **Regulatory background**
- ✓ **Comparability analysis on Stability studies from accelerated/stressed conditions**
- ✓ **Comparability analysis on Release data from routine manufacturing**

AGENDA

- ✓ **Regulatory background**
- ✓ **Comparability analysis on Stability studies from accelerated/stressed conditions**
- ✓ **Comparability analysis on Release data from routine manufacturing**

Regulatory background - ICH Q5E*

PURPOSE

- Comparing post-change product to pre-change product following manufacturing process changes

When considering the comparability of products, the manufacturer should evaluate, for example:

- **The need for stability data**, including those generated from accelerated or stress conditions, to **provide insight into potential product differences** in the degradation pathways of the product and, hence, potential differences in product-related substances and product-related impurities;
- **Accelerated and stress stability studies are often useful tools to establish degradation profiles** and provide a further **direct comparison of pre-change and post-change product.**

*Guidance for Industry Q5E Comparability of Biotechnology/Biological Products Subject to Changes in Their Manufacturing Process (June 2005)

Regulatory background – EMA reflection paper*

PURPOSE

- Comparing post-change product to pre-change product following manufacturing process changes

Practical considerations for comparability of products:

- In practice, **comparability ranges are frequently established based on a statistical interval**, e.g. the min-max range or a tolerance interval calculated from characterization data of the reference product.
- **comparison of single batch data to a min-max range might be suitable in the context of batch-release**
- A tolerance interval (TI) is usually computed to estimate a data range by which a specified proportion (e.g. the central 90%) of the units from the underlying population is assumed to be covered with a pre-specified degree of confidence (e.g. 95%) … **all test batches of the sample fall within the 90%/95% TI computed from the reference batches**

*Reflection paper on statistical methodology for the 4 comparative assessment of quality attributes in drug 5 development (March 2017)

AGENDA

- ✓ Regulatory background
- ✓ **Comparability analysis on Stability studies from accelerated/stressed conditions**
- ✓ Comparability analysis on Release data from routine manufacturing

Comparability pre-/post- change for stability data

General context

Context:

- Process change (ex: new improved process, new site)
- 6 manufactured batches (3 pre- and 3 post- change), consecutive batches are usually chosen for each process
- DS/DS or DP/DP comparability
- Only the stability indicating methods are selected
- Stability at accelerated/stressed condition is performed, duration is chosen so it is representative of the total degradation that will occur at the intended storage condition for the shelf-life period
- **Comparability protocol:** degradation rate between pre- and post-change batches at accelerated/stressed condition are similar

For major changes in order to reduce the analytical variability:

- Batches are run on stability in parallel
- The stability samples are analyzed in side-by-side analysis (in the same analytical sequence) when feasible.

Comparability pre-/post- change for stability data

Types of comparability

The following types of comparability are done:

- **For decreasing or increasing attributes for which sufficient quantifiable data are available (at least 3 time points with values above LOQ by batch):** Comparison of slopes and intercepts among processes by mixed effects ANOVA: test for difference
- **For increasing attributes with insufficient quantifiable data (less than 3 data points with values above LOQ for at least one batch):** Comparison of probability of increased risk of Out Of Specification (OOS) values (between original and new process) and comparison of ranges of values

Comparability pre-/post- change for stability data

Decreasing or increasing attributes with sufficient quantifiable data

For decreasing or increasing attributes for which sufficient quantifiable data are available (at least 3 time points with values above LOQ by batch):

- Estimate degradation rates for each process via a mixed effects ANOVA model
- Use “process” (2 levels: pre- and post- changes process) as a fixed effect, “batch within process” as a random effect, and “time” as a covariate.
- Example of SAS code:

```
proc mixed data;
    class batch Process ;
    model response = time   Process   time*Process/s;
    random batch(Process)    time*batch(Process)/s;
    run;
```

- A test for slopes and intercepts between process is conducted

Comparability pre-/post- change for stability data

Decreasing or increasing attributes with sufficient quantifiable data

To determine the poolability of different processes the following tests are performed:

1. Test for equality of slopes ("time*Process")
2. Test for equality of intercepts ("Process")

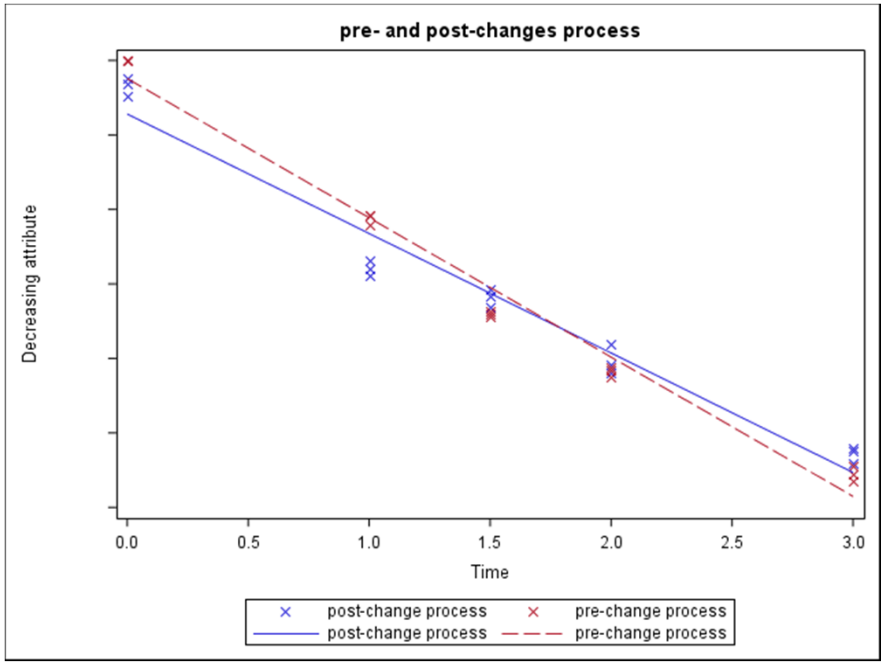
Statistical analysis is performed at the significance level of 5% (alpha=0.05)

Based on these hypothesis, three different models can be proposed:

- **Model 1:** Separate slopes and separate intercepts: Degradation profiles of the tested processes are not homogeneous. They differ in their degradation rate.
- **Model 2:** Common slope but different intercepts: Degradation profiles of the tested processes behave the same in their degradation rate but they differ by an offset.
- **Model 3:** Single common regression model: Degradation profiles of the tested processes have a common slope and common intercept. Processes have the same degradation rate.

Comparability pre-/post- change for stability data

Decreasing attributes – case study 1 (Model 1)



Parameter	p-value	Conclusion
time*process	**0.0259**	**<0.05**

Model 1: Separate slopes and separate intercepts

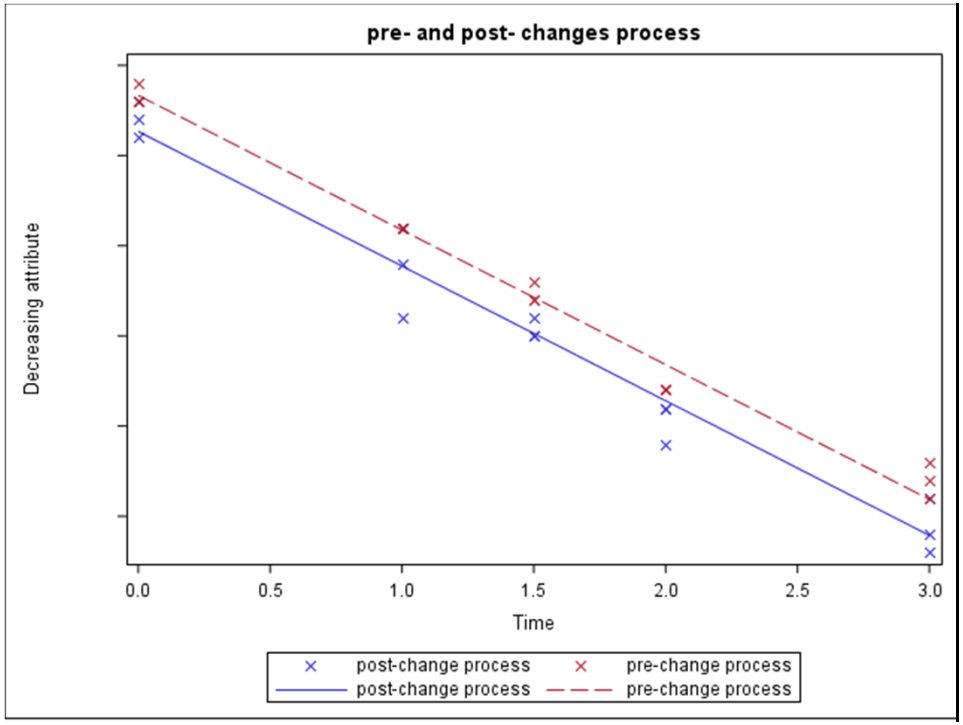
Estimated difference between total degradation at 3 months: 0.79%

This is less than the analytical variability of 1.5% => degradation rates are considered the same

A comparison of degradation slopes at the intended storage condition was also performed with the same methodology and confirmed that slopes are comparable (p-value >0.05)

Comparability pre-/post- change for stability data

Decreasing attributes – case study 2 (Model 2)



Parameter	p-value	Conclusion
time*process	**0.4050**	**>0.05**
process	**0.0459**	**<0.05**

Model 2: Common slopes and separate intercepts

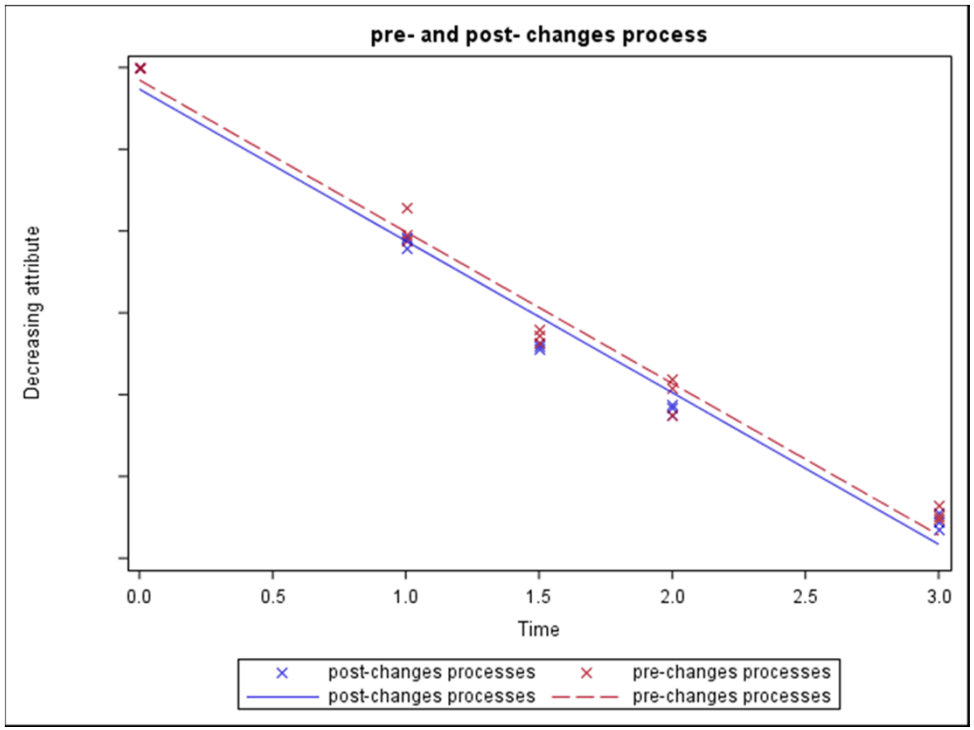
=> degradation rates are considered the same

Estimated difference between intercepts: 0.20%

This is within the expected variability between batches
=> intercepts are considered the same

Comparability pre-/post- change for stability data

Decreasing attributes – case study 3 (Model 3)



Parameter	p-value	Conclusion
time*process	**0.7576**	**>0.05**
process	**0.2567**	**>0.05**

P-value >0.05

Model 3: Single common regression model

=> degradation rates and intercepts are considered the same

Comparability pre-/post- change for stability data

Increasing attribute with values <LOQ

- The **mixed model approach cannot be applied** for increasing attributes for which not sufficient quantifiable data are available (regression cannot be estimated)
- **The comparability between processes is done by comparing:**
  - the number of OOS results versus the number of results within specification at each time point.
  - the range of values from post-change batches with the range of values from pre-change batches

- **If the range of values** from post-change batches is **within or equal** to the range of values from pre-change batches, **process are considered comparable**
- **If the range of values** from post-change batches is **larger than** the range of values from pre-change batches, a **comparison of the observed difference with the analytical variability is made**

Comparability pre-/post- change for stability data

Increasing attribute with values <LOQ

Example: For an increasing attribute, with LOQ=0.4% and Specification=1% the following values were observed:

Time point (Months)	Results number			
	Pre-change process		Post-change process	
	In specification	OOS	In specification	OOS
0	3 (<LOQ)	0	3 (<LOQ)	0
1	3 (<LOQ)	0	3 (<LOQ)	0
1.5	3 (<LOQ)	0	3 (<LOQ)	0
2	3	0	3	0
3	0	3	0	3

- Up to 1.5M, all values are below LOQ for both pre- and post-change- process
- At 2M, all values are in specification and observed values for pre-change process are 0.8%, 0.8%, 0.8% and for post-change process are 0.7%, 0.8%, 0.8%.
- At 3M, all values are OOS and observed values for pre-change process are 1.2%, 1.2%, 1.2% and for post-change process are 1.1%, 1.2%, 1.2%.

➢ **Process are considered comparable with respect to this attribute**

AGENDA

- ✓ Regulatory background
- ✓ Comparability analysis on Stability studies from accelerated/stressed conditions
- ✓ **Comparability analysis on Release data from routine manufacturing**

Comparability pre-/post- change for release data

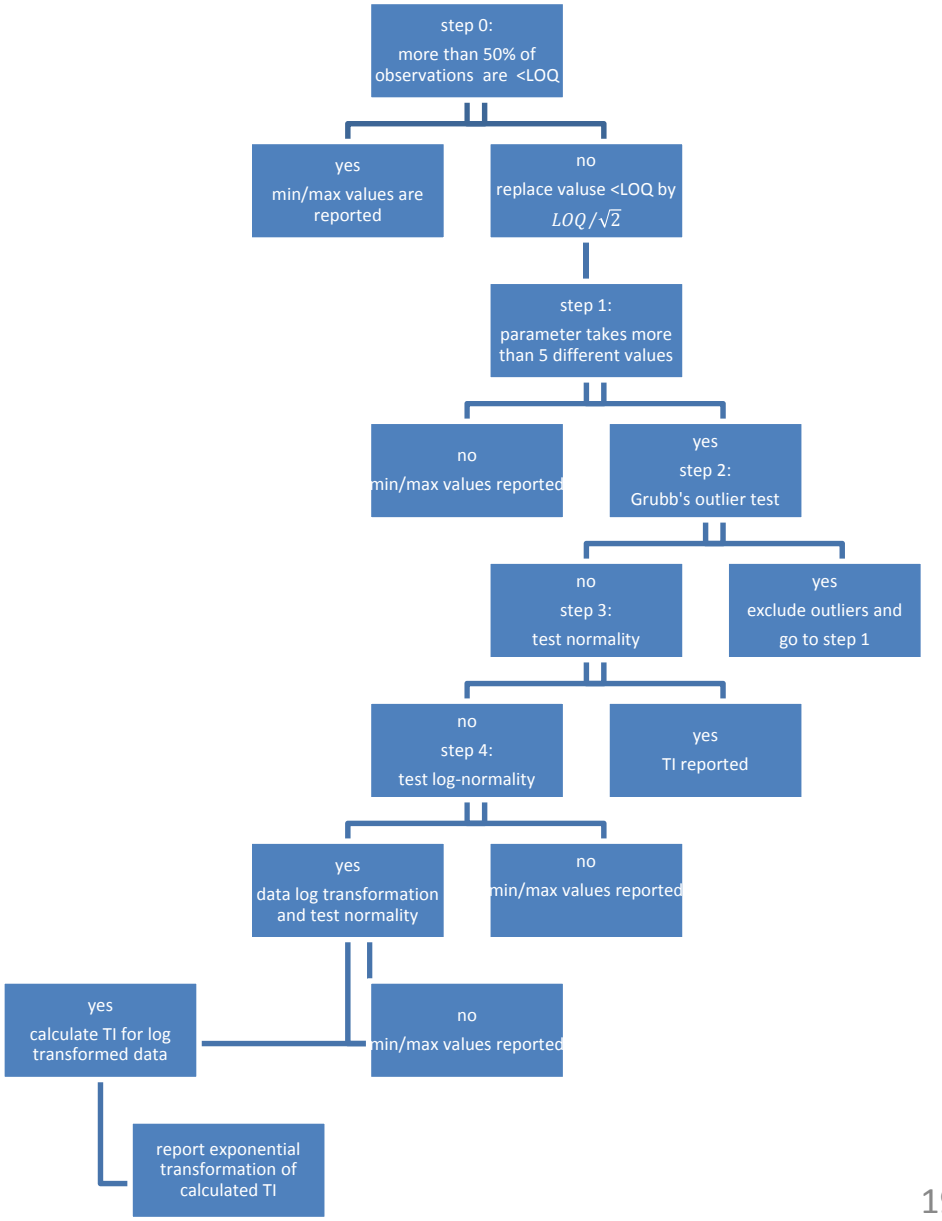
Context:

- Process change (ex: new improved process, new site)
- 3 manufactured batches post- change (validation batches)
- Large pool of historical data (including clinical batches) for pre-change process
- Attributes are divided into three categories:
    1. qualitative attributes
    2. quantitative attributes with values below the limit of quantification (LOQ)
    3. quantitative attributes with values above LOQ.
- Only the quantitative attributes are discussed in this presentation
- Historical ranges are determined as:
    1. **min:** minimum value observed (or NA); **max:** maximum value observed (or <LOQ).
    2. **Tolerance Intervals (TI)** covering 99% of the population with a 95% confidence level.

- **Comparability protocol:** release values for batches post-change are within the historical range for pre-change batches
- **Post change monitoring:** Release values for post-change process are monitored through Continued Process Verification plan, through product lifecycle

Comparability pre-/post- change for release data

- **If more than 50% of the observations are below the limit of quantification** (LOQ), no statistical analysis is performed; limits are fixed as: min/max.
- **If less than 50% of the observations are below the LOQ**, individual data reported as below the LOQ are replaced by LOQ 1/(√2) .
- **If less than 6 different values are reported for an attribute**, limits are fixed as min/max.
- **If at least 6 different values are reported for an attributed**, the normality is verified using the p-value of Shapiro-Wilk test or Kolmogorov's D test. If data are normally distributed limits are fixed as 99%95% TI. If data are neither normally distributed nor log-normally distributed, limits are fixed as min/max.
- Outliers are excluded from TI computation (Grubbs test to detect outliers).



Comparability pre-/post- change for release data

Examples

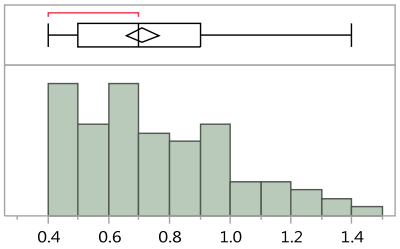
Example 1: Continuous parameter with more than 50% of observations <LOQ

=> Historical limits: min/max= <LOQ/0.5 **post-changes values:** <LOQ, <LOQ, <LOQ

Example 2: Continuous parameter with 20% of observations <LOQ ;

- Data below LOQ are replaced by $LOQ\,\frac{1}{\sqrt{2}}$,
- Data are not normally nor log-normally distributed

=> Historical limits: min/max= <LOQ/1.4 **post-changes values:** 0.6, 0.8, 0.9

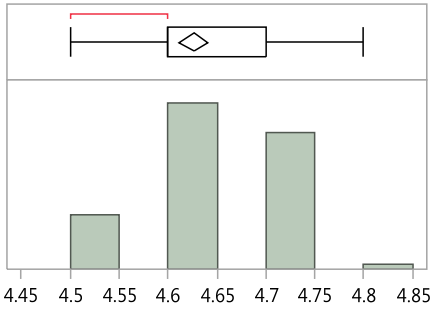


Comparability pre-/post- change for release data

Examples

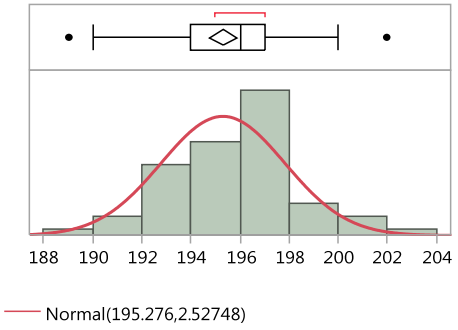
Example 3: Continuous parameter with less than 6 different values (ex: pH)

=> Historical limits: min/max= 4.5/4.8 **post-changes values:** 4.6, 4.6, 4.7



Example 4: Continuous parameter with at least 6 different values

=> Historical limits: 99%95%TI: [188, 203] **post-changes values:** 197, 199, 200



TAKE AWAY MESSAGE

- Stability data should be considered for process pre-/post-change comparison.
- Evaluation of stability data at recommended storage conditions would require entire expiry period, for a good estimation of slope and a meaningful comparison. Instead, evaluate stability data at accelerated/stressed condition.
- Accelerated conditions can provide a direct comparison of pre- and post-change product that might not be apparent at lot release or recommended storage.
- For batch release comparability, compute historical ranges (min/max or TI depending on the available data) and compare with release values from validation batches.

Questions…