



Comparability in Accelerated or Stressed Stability Studies Using a Quality Range

Richard K. Burdick

A statistical procedure based on a quality range is proposed for demonstrating comparability between pre-change and post-change process slopes in an accelerated or stressed stability study.

To assess the impact of process changes, stability data (e.g., measurements of drug product over time) at recommended storage conditions are often required. Pre- and post-change degradation rates (slope over time) do not need to be identical, but do need to be highly similar, with scientific justification that any observed differences will not impact safety or efficacy. Because slopes at the recommended storage conditions are often quite small, the product is also exposed to non-recommended storage conditions (e.g., higher temperature)

Richard K. Burdick is principal of Burdick Statistical Consulting and Professor Emeritus of Statistics at Arizona State University.

to quickly accelerate degradation. In comparability studies, it is necessary to demonstrate that the degradation rates (slopes) of the pre-change and post-change processes are comparable. Burdick and Sidor proposed a statistical test of equivalence that can be applied in these situations (1). However, it might be the case that practically sized experiments may lack sufficient power to perform the equivalence test. An alternative approach is proposed based on the quality range test recommended by FDA in their draft guidance for comparing biosimilars (2). This approach uses a visual interpretation of study results based on a heuristic rule for claiming process comparability.

Example

A typical study design used in an accelerated experiment requires $L=3$ lots from both the pre-change and post-change processes. There are typically $T=3$ time points measured on each of the six lots with one measurement per lot at each time point. **Figure 1** displays plots of two data sets with $L=3$ lots, $T=3$ time points ($t_1=0$, $t_2=14$, and $t_3=28$ days), and one measurement at each time point (a total of nine values for each process). The response variable is cation exchange (CEX) Main (%). The lines shown in the figure are the least square lines of fit for each of the six lots. The complete data set is provided in **Table 1**.

Of particular interest is the comparison of the line slopes for the two processes. Slopes differ within each process due to analytical error. Differences between processes indicate a difference in degradation rate between the two processes and should be investigated for impact on safety or efficacy.

A visual representation that allows one to focus exclusively on slopes is provided in **Figure 2**. Here suitable constants have been added to each y-intercept so that they are all equal to the same (arbitrary) value.

Note in **Figure 2** that the slopes from the two processes in Data Set 1 overlap the same region and there is no clear delineation between the two processes. In contrast, all three slopes of the post-change process are steeper than those of the pre-change process in Data Set 2, and there is no overlap. One might declare comparability in Data Set 1; whereas with Data Set 2, it appears the processes may not be comparable.

Quality range test for comparing slopes

The quality range test provides a statistical heuristic based on a visual comparison of the slope values of the pre-change and post-change lots. The quality range test was introduced by FDA in their 2019 guidance document for analytical comparison of biosimilar products (2). One can perform the quality range test for the accelerated study by following these steps:

Figure 1. Plot of data from accelerated study.

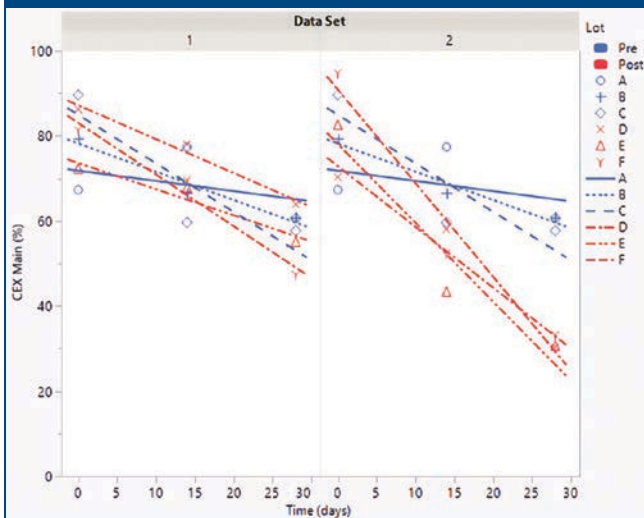
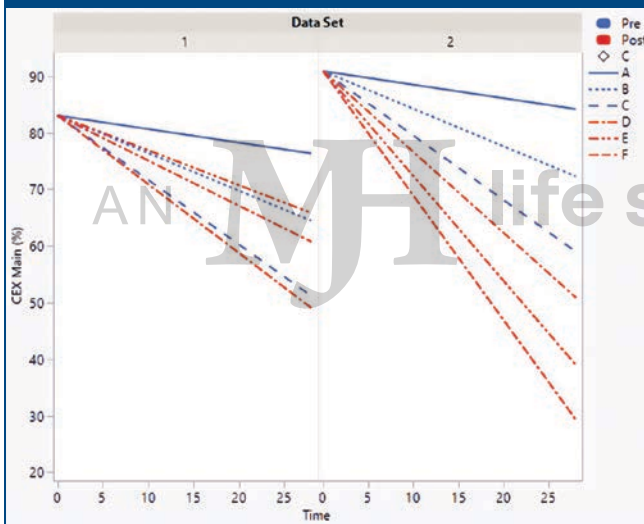


Figure 2. Degradation lines plotted with equal intercepts.



1. Compute the average of the L pre-change lot slopes. Denote this statistic as b_{pre} . The L=3 pre-change slopes are -0.239 (A), -0.664 (B), and -1.139 (C) in both data sets. The average of these three lots is $b_{pre} = -0.681$.

2. Establish a value for the analytical standard deviation. This value can be obtained from many external sources including assay control data, validation/qualification studies, and long-term stability measurements. If no external source is available, the analytical standard deviation can be estimated using the pooled

error from each of the individual lot regression models (both pre- and post-change). Denote this standard deviation as $\sigma_{Analytical}$. For this example, $\sigma_{Analytical} = 10$.

3. Compute the lower and upper quality range limits using Equation 1:

$$LQL = b_{pre} - C \times \frac{\sigma_{Analytical}}{\sqrt{SST}}$$

$$UQL = b_{pre} + C \times \frac{\sigma_{Analytical}}{\sqrt{SST}}$$

(1) [Eq.1]

Table I. Example data sets.

Group	Lot	Time (days)	CEX Main (%)	Data Set
Pre	A	0	67.3	1 and 2
Pre	A	14	77.4	1 and 2
Pre	A	28	60.6	1 and 2
Pre	B	0	79.4	1 and 2
Pre	B	14	66.5	1 and 2
Pre	B	28	60.8	1 and 2
Pre	C	0	89.6	1 and 2
Pre	C	14	59.6	1 and 2
Pre	C	28	57.7	1 and 2
Post	D	0	86.2	1
Post	D	14	78	1
Post	D	28	63.9	1
Post	E	0	72.2	1
Post	E	14	67.8	1
Post	E	28	55	1
Post	F	0	81.3	1
Post	F	14	69.5	1
Post	F	28	47.3	1
Post	D	0	70.3	2
Post	D	14	58	2
Post	D	28	30.3	2
Post	E	0	82.5	2
Post	E	14	43.3	2
Post	E	28	30.6	2
Post	F	0	94.7	2
Post	F	14	52.4	2
Post	F	28	33.1	2

where, $SST = \sum_{i=1}^T (t_i - \bar{t})^2$, $\bar{t} = \frac{\sum_{i=1}^T t_i}{T}$, and C is a suitably chosen constant. (Considerations for selecting C are discussed in the next section.)

Here $SST = (0-14)^2 + (14-14)^2 + (28-14)^2 = 392$.

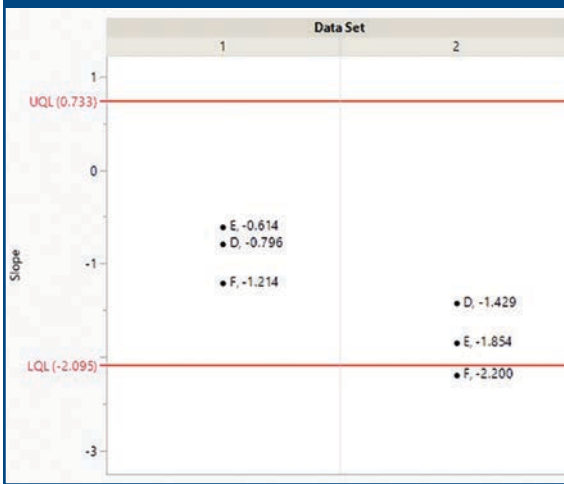
Using Equation 1 with C=2.8, the resulting calculation is

$$LQL = -0.681 - 2.8 \times \frac{10}{\sqrt{392}} = -2.095$$

$$UQL = -0.681 + 2.8 \times \frac{10}{\sqrt{392}} = 0.733$$

(2) [Eq. 2]

4. Comparability is declared if all L post-change lot slopes fall

Figure 3. Quality range and slopes.**Table II.** Probabilities of passing for values of C and T (L=3).

Design	C	T	Slopes equal	Slopes "Clearly different"
1	2.8	3	0.95	0.13
2	2.6	3	0.93	0.09
3	3.0	3	0.97	0.18
4	2.8	4*	0.95	0.13

*Timepoints are 0, 9, 18, 27 days with SST=405.

within the quality range based on the pre-change process.

Figure 3 provides the quality range and post-change process slopes for both data sets. The numbers shown in the body of the figure are the values of the post-change slopes.

Note that all post-change slopes fall within the quality range for Data Set 1, but that one of the post-change slopes (F) falls outside the range in Data Set 2. Thus, one might conclude comparability for Data Set 1 but not Data Set 2, matching the visual conclusion in **Figure 2**.

Additional considerations

The quality range approach assumes that the measured values of the response are independent. It may happen that a lab will "batch" all results of the stability study and analyze in a single assay. This will violate the independence assumption if the analytical procedure manifests any assay-to-assay variability. If such is the case, then

$\sigma_{\text{Analytical}}$ should represent only the within assay variance.

The selection of C should provide a high probability of passing when the two process slopes are equal and a low probability of passing when the two process slopes are "clearly different". A workable definition of "clearly different" is when the true difference between the pre-change and post-change slopes is at least three standard errors of the slope (i.e., at least $(3\sigma_{\text{Analytical}})/\sqrt{\text{SST}}$). Visually, a difference of this magnitude would manifest in a plot like Data Set 2 in **Figure 2**

where there is no overlap of the individual lot slopes for the two processes. The value of C=2.8 selected in this example provides a 95% probability of passing when the two process slopes are equal. Conversely, if the process slopes are "clearly different", the probability of

passing is only 13%. **Table II** reports probabilities for other values of C and T. These values are easily obtained by simulation.

The fact that Designs 1 and 4 produce the same results demonstrates that increasing T beyond 3 has little if any impact.

Conclusion

Figure 1 presents data from a typically sized accelerated stability study. Conversion to the plot in **Figure 2** allows one to better visualize any differences that might exist between slopes of the pre-change and post-change processes. An additional plot provided in **Figure 3** allows use of a heuristic criterion for determining comparability of the two processes. JMP script that produces all three figures is available by request from the author.

References

1. Bürdick, R.K.; Sidor, L. Establishment of an Equivalence Acceptance Criterion for Accelerated Stability Studies. *Journal of Biopharmaceutical Statistics*, **2013** 23: 730-743, DOI: 10.1080/10543406.2013.789891
2. FDA. *Development of Therapeutic Protein Biosimilars: Comparative Analytical Assessment and Other Quality-Related Considerations, Draft Guidance* (CDER, CDER, 2019). **PT**

Drug Digest: Analytics and Assays Strive to Keep Pace with Innovative Modalities

With the evolution of biologic modalities in the drug development pipeline, the biopharma industry understands that it must keep pace with the development and use of appropriate analytics and assays for molecule characterization to determine product purity, safety, and efficacy, and even early on in the R&D process, where analytics are key in screening multitudes of molecules to determine viable candidates to move into the drug development phases.

This episode of *Drug Digest* explores new and emerging technologies, such as mass photometry, machine learning, and AI, and their role in the analytical space for biologic modalities.

Scan the QR code to watch this episode.

—The Editors of
Pharmaceutical Technology®

