AMGEN®

Pioneering science delivers vital medicines™

Establishing Equivalence Acceptance Criteria for Accelerated Stability Studies

Leslie Sidor, Rick Burdick and Camilla Santos

Amgen, Inc

36th Annual MBSW, May 22, 2013

Agenda

- Use of accelerated/stressed stability
- Determination of acceptance criterion
- Example
- Conclusions
- References

Product comparability is driven by ICH Q5E[1]

- Risk based evaluation of magnitude of change →potential impact of change on product quality attributes
  - Scale change
  - Site change
  - Cell line change
- Pre- and post-change products need to be highly similar, with scientific justification that any observed differences will not impact safety or efficacy.
- Comparability studies should be designed with pre-defined acceptance criteria.

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

Use of accelerated stability is encouraged for biologics

- Evaluation of recommended storage conditions has limited value
  - Minimal degradation
  - Best slope estimate requires entire expiry period
- Biologics do not follow Arrhenius behavior → cannot link to performance at recommended storage conditions
- Accelerated temperatures can provide a direct comparison of pre and post change product that might not be apparent at lot release or recommended storage

Accelerated stability is typically thermal stress

Comparing slopes from accelerated stability data has multiple options

Issue	Visual Assessment	Difference Test (p-value)	Equivalence of Slopes
Objective assessment	No	Yes	Yes
Patient risk controlled	No	No	Yes – Fixed at 5 %
Manufacturer risk controlled	No	Yes – Fixed at 5 %	Yes – if sample size sufficient
Endorsed in the literature	Possibly	No	Yes
Test can prove equivalence	No	No	Yes
Science is built into acceptance criteria	No	No	Yes

Use of equivalence makes the most sense for comparing slopes under accelerated stability conditions

How does one define an EAC for use in an Equivalence Test?

- To be meaningful, the EAC must be defined by a subject matter expert (SME) prior to the experiment.
- Schuirmann (1987) states the specification of EAC "is made by experts in the fields of biopharmaceutics and medicine (*not by the statistician*!)"
- However, it is the responsibility of the statistician to help the SME with this selection and present options that are easily understood by the SME.
  - Wellek (2010) argues that the statistician must "provide the experimental or clinical researcher with a range of options sufficiently large for allowing him to cover the question he really wants to answer by means of his data".

Problem with accelerated stability data → stability specifications cannot be used to provide a definition of unacceptability

EAC With No Definition of "Acceptable" From the SME

- Hauck et al. (2008) refer to equivalence criteria based on what is either "acceptable" or "unusual".
- Examples for non-stability data when a definition of "acceptable" exists are found in USP <1010> Appendix E and Chatfield and Borman (2009).
- We propose a visual capability approach that identifies what would seem "unusual" to an SME.
- This approach is recommended for the accelerated stability study where no meaningful specifications exist.
- The approach we advocate is based on a visualization of change in terms of the effect size.

Utilize effect size to describe an "unusual" shift → definition is now based on what is "expected"

Accelerated Stability Model: Random Intercept Model

- **A statistical model that represents measurements of a quality attribute from the historical process is**

$$Y_{Hij} = \mu_H + L_i + \beta_H t_j + E_{ij}$$

$$i = 1,\ldots,n_H;\ j = 1,\ldots,T$$

L_i independent normal random variables with mean 0 and variance σ_L^2

E_{ij} independent normal random variables with mean 0 and variance σ_E^2

L_i and E_{ij} are independent

- **For the new process, assume the same error structure with the following hypothesis**

$$H_0 : \left|\beta_H - \beta_N\right| \geq EAC$$

$$H_1 : -EAC < \beta_H - \beta_N < EAC$$

When subject matter expertise cannot define an EAC, what are the options?

Consider using an effect size and the distribution of the historical slopes

$$ES = \frac{\left|\beta_H - \beta_N\right|}{\sqrt{Var\left(\hat{\beta}_{Hi}\right)}}$$

$$Var\left(\hat{\beta}_{Hi}\right) = \frac{\sigma_E^2}{SST} \text{ where } SST = \sum_{j=1}^{T}\left(t_j - \bar{t}\right)^2$$

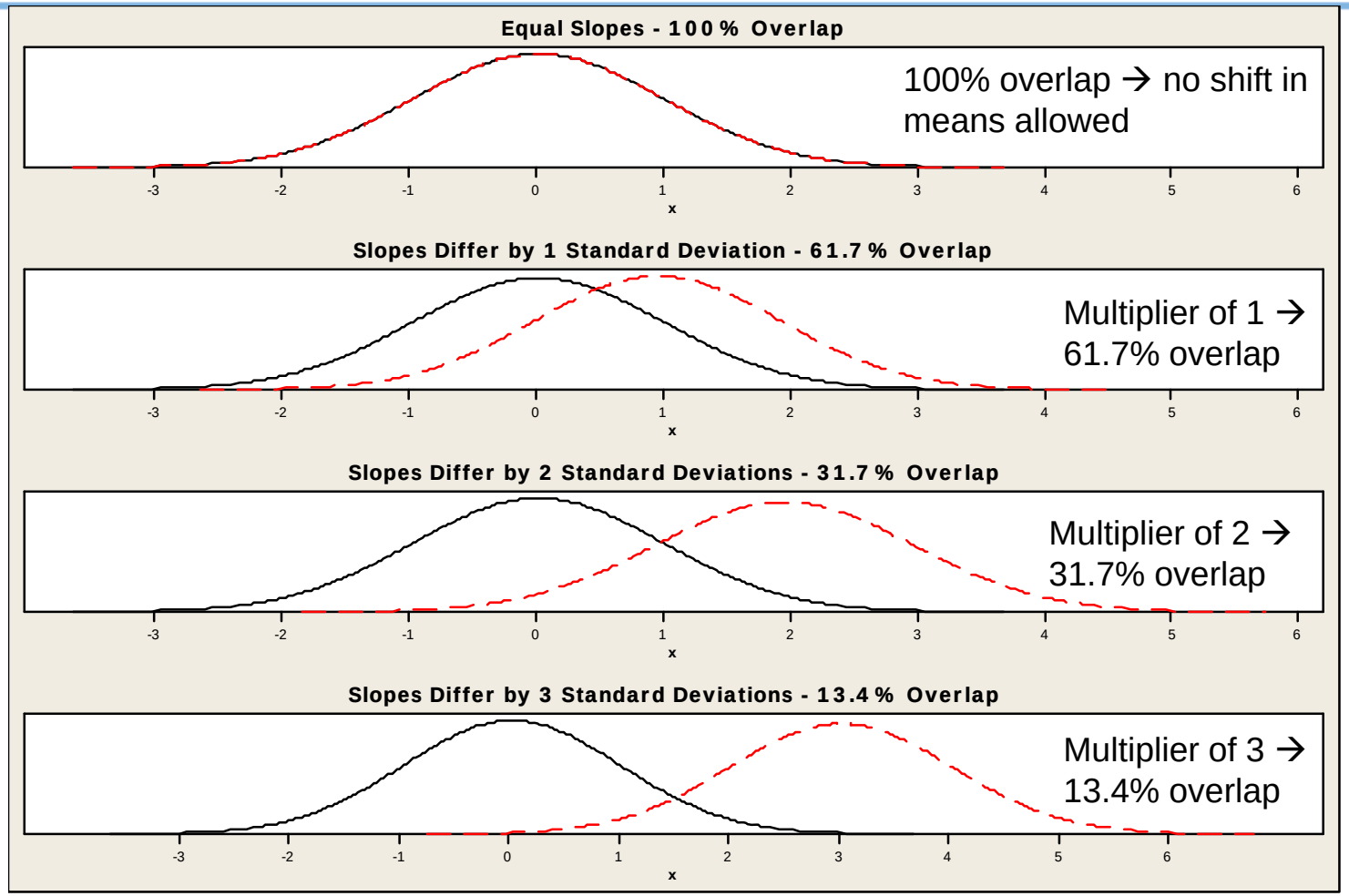
Now the EAC becomes

$$EAC = ES \times \sqrt{Var\left(\hat{\beta}_{Hi}\right)}$$

$$EAC = ES \times \sqrt{\frac{\sigma_E^2}{SST}}$$

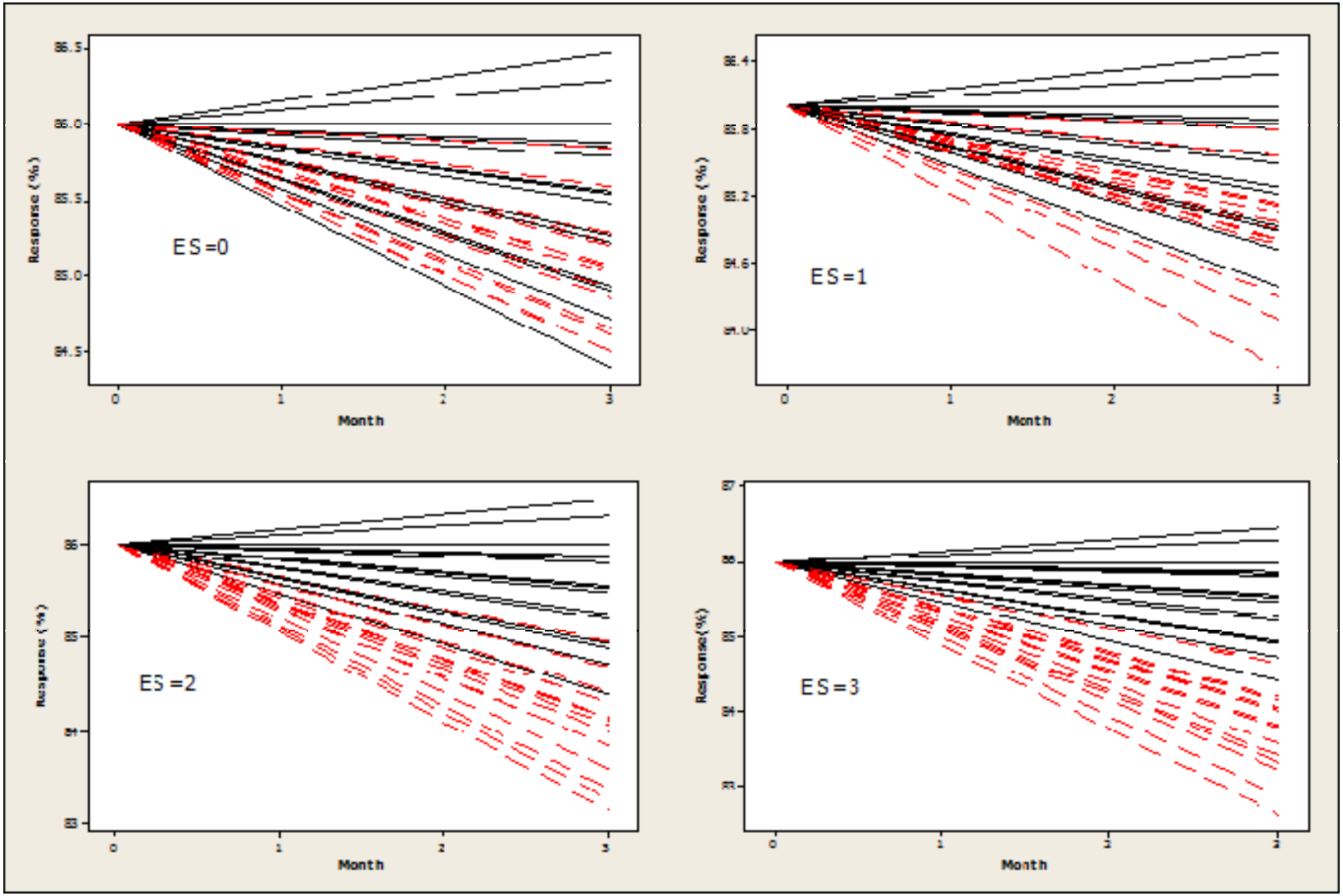
- Note that the EAC is a function of the analytical method error and the stability design (SST)

What is an appropriate effect size?



SME can consider the overlap of the historical and new slope distributions to identify an EAC

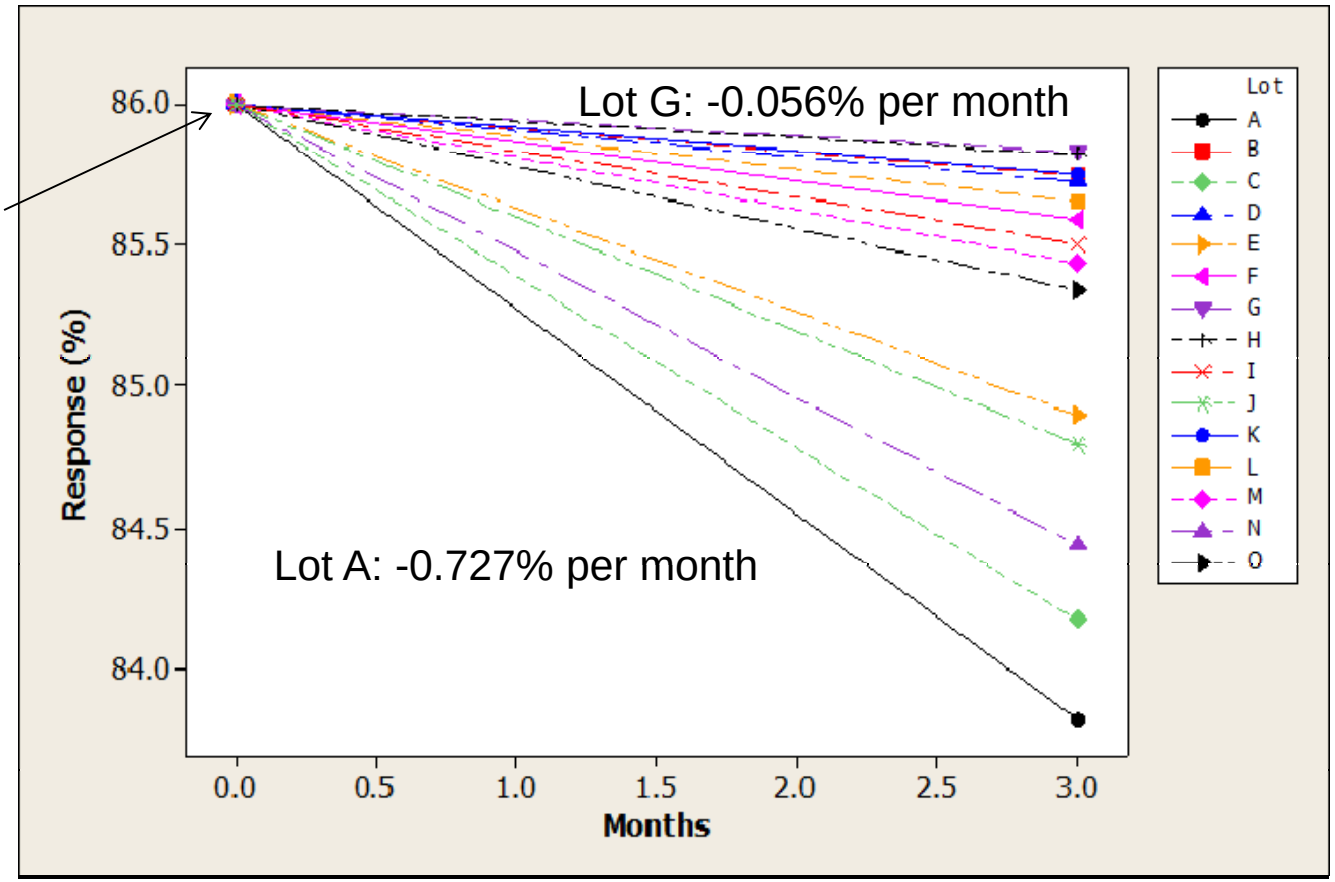
A more effective set of plots to help select a reasonable Effect Size for stability data



Example: % Purity, Product stored at 37ºC for 3 months

- Historical data set: 15 lots
- Purity measured at 0, 1, 2 and 3 months (SST=5)

Lines adjusted to have the same y-intercept



Results from fitting the random intercept model for the historical data

Parameter	Estimate
β_H	-0.255
σ_L^2	0.428
σ_E^2	0.200
95% Upper bound on σ_E^2	0.295
SST	5

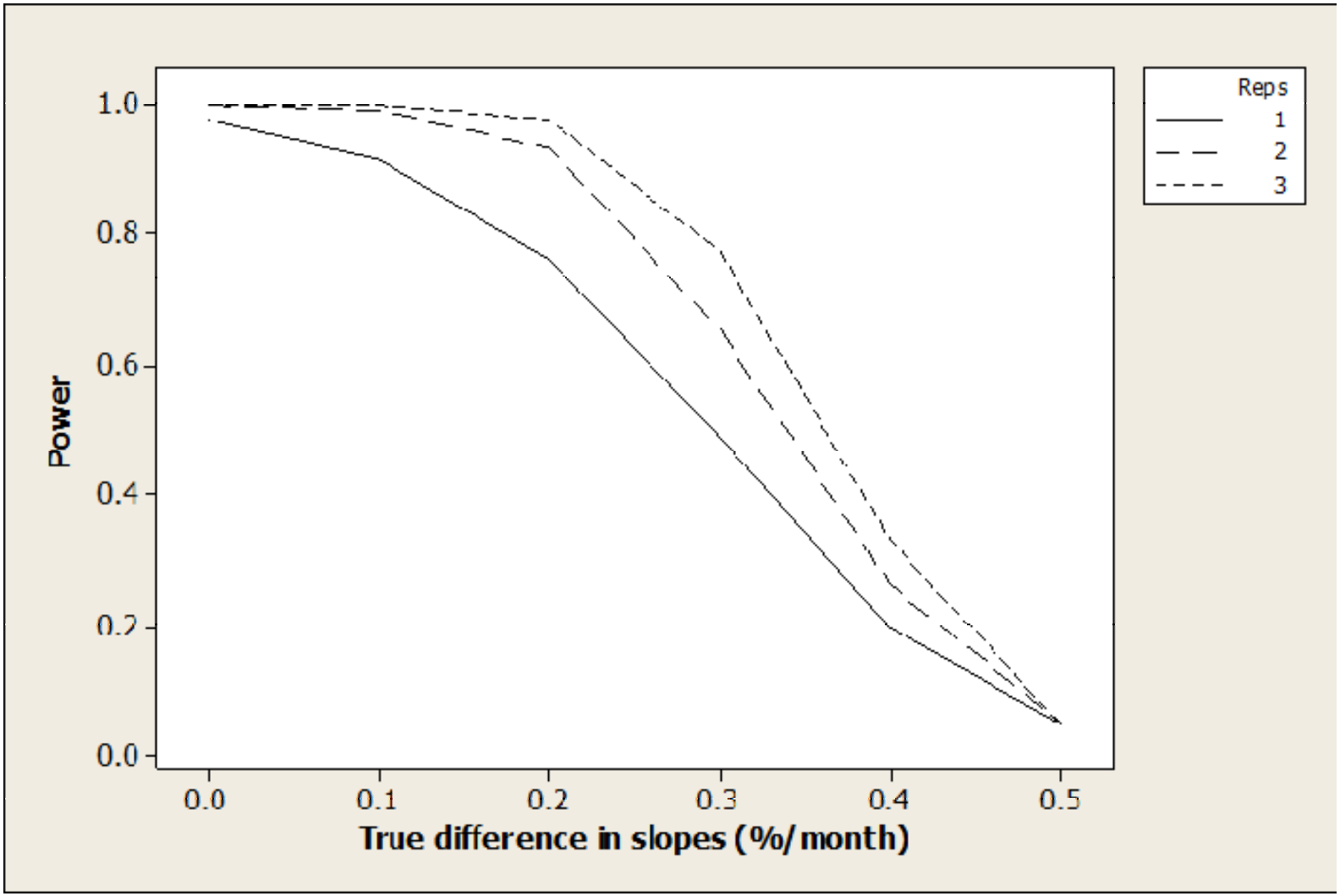
$$EAC = 2 \times \sqrt{\frac{0.295}{5}} = 0.49 = 0.5\ \%\ \text{per month}$$

The upper bound is used to accommodate sampling error

Control of Type 2 Error rate

- Evaluate power to determine the required sample size needed to ensure producer risk is acceptable
- A typical stability study consists of a comparison of many historical lots against a very few new process lots
  - Typical values might be 15 lots for the historical process and 3 lots for the new process
  - Even with only 3 new process lots, power can be increased under the random intercept model by performing replicate stability studies for each new lot
  - Replication is most effective at t=0 and the final time point

Power Curves for Replicated Studies With 3 New Lots and 15 Historical Lots



TOST Calculations

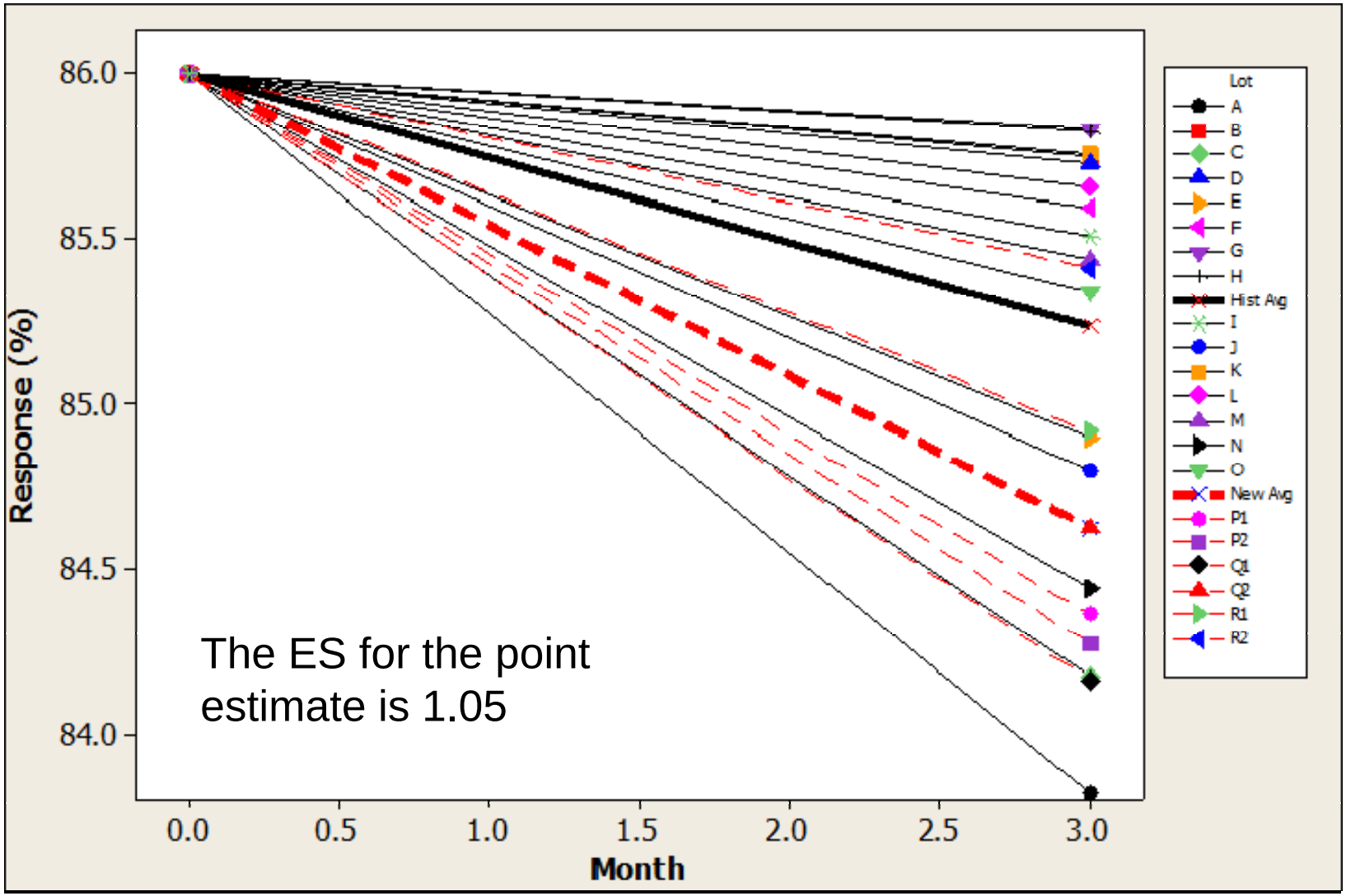
- **Lot is nested within Process, and Stability Study nested within Lot**

σ^2_{Study}	0
σ^2_L	0.379
σ^2_E	0.190

	Estimate	Lower 95% One-sided Bound	Upper 95% One-sided Bound
Historical Slope	-0.255	-0.339	-0.171
New Slope	-0.459	-0.592	-0.326
Difference (H-N)	0.204	0.047	0.361

90% two-sided interval on the difference: 0.047%/month to 0.361%/month satisfies EAC=0.5%/month

Slope Estimates for Each Process
(Red dashed is new and Black solid historical)



Concluding Remarks

- A capability approach using a visualization of ES aids the SME in determining an appropriate EAC when no definition of “acceptable” is provided.
- Power calculations should follow EAC determination in order to determine producer’s risk.

References

- Burdick, R. and Sidor, L. (2013). Establishment of an Equivalence Acceptance Criteria for Accelerated Stability Studies. *Journal of Biopharmaceutical Statistics*, Forthcoming
- Chambers, D., Kelly, G. Limentani, G., Lister, A., Lung, R., Warner, E. (2005). Analytical method equivalency: An acceptable analytical practice. *Pharmaceutical Technology*, 29, 64-80.
- Chatfield, M., Borman, P. (2009). Acceptance criteria for method equivalency assessments. *Analytical Chemistry*, 81, 9841-9848.
- Hauck, W. W., Abernethy, D. R., Williams, R. L. (2008). Metrologic approaches to setting acceptance criteria: Unacceptable and unusual characteristics. *Journal of Pharmaceutical and Biomedical Analysis*, 48, 1042-1045.
- ICH Q5E Comparability of Biotechnological/Biological Products Subject to Change in Their Manufacturing Process; June 2005
- Limentani, G., Ringo, M., Ye, F., Bergquist, M., McSorley, E. (2005). Beyond the t-test: Statistical equivalence testing. *Analytical Chemistry*, 77, Issue 11, 221A-226A.
- Schofield, T. (2009). Maintenance of vaccine stability through annual stability and comparability studies. *Biologicals*, 37, 397-402.
- Schuirmann, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, 15, 657-680.
- Wellek, S. (2010). Testing Statistical Hypotheses of Equivalence and Noninferiority, Second Edition. Chapman & Hall/CRC

Back up

Definition of EAC for Accelerated Stability

- **Schofield (2009) has considered "acceptable" criteria for the stability problem at recommended conditions and proposes an equivalence test based on a difference in slopes and an acceptability requirement at the end of shelf life.**
- **While it might be possible to establish equivalence criteria by using Arrhenius kinetics to link acceptable degradation rates at accelerated conditions to product specifications at recommended conditions, such kinetics rarely apply to biological product degradation mechanisms, and this is not considered a generally useful approach.**

Equivalence Test: Difference in Slopes vs. Ratio of Slopes

Equivalence Test	Attributes
Difference in Slopes	•Describes change difference in the quality attribute over a fixed period of time •Difference in slopes has a meaningful unit of measure for SME
Ratio of Slopes	•A ratio of slopes has no unit of measure or meaningful interpretation •Ratio of slopes is not always consistent with visual representation •Cannot be defined if slopes close to zero have different signs

It is recommended to use a difference in slopes as the equivalence test for product comparability

Ratio of Slopes Not Always Consistent with Visual Representation