Perspectives on Pooling as Described in the ICH Q1E Guidance

Stan Altan[1], Areti Manola[1], Jyh-Ming Shoung[1],Yan Shen[1]
[1]Janssen Research & Development, LLC, 1000 Route 202S, Raritan, NJ 08869

Abstract

The June 2004 ICH Q1E guidance lays out a statistical pooling strategy applicable to stability studies carried out according to the principles given in the Q1A(R2) guidance. The strategy assumes a completely fixed batch-specific model to describe the concentration of the active pharmaceutical ingredient in relation to time, in which hypothesis testing for poolability of slopes across batches is followed by poolability of intercepts. The criterion for poolability is set at p-value=0.25. The approach is extended when pooling across fixed factors, such as package or strength. These recommendations tilt toward statistical convenience with little regard to chemical or process engineering considerations. Moreover, more recent approaches such as mixed modeling and Bayesian hierarchical modeling tools are readily available which are not discussed in the guidance. This presentation proposes a mixed modeling framework and reshapes the pooling recommendations given in the Q1E guidance based on scientific and empirical claims.

Key Words: Stability modeling, ICH Q1E Guidance, Mixed Modeling

1. Introduction

Stability studies are an important part of drug development, falling within the general purview of nonclinical studies belonging to the Chemistry, Manufacturing and Control (CM&C) aspects. The CM&C section constitutes a major part of the New Drug Application (NDA) process. Typically, a stability protocol is written which lays out the design of the study, defining the number of batches of drug product and fixed factors under study, such as strength and package. The essential objective of a stability study is to estimate rates of change of the active pharmaceutical ingredient (API) and other properties of the drug substance or packaged drug product in relation to storage time under various conditions related to the ICH defined stability zones. The statistical model then leads to a shelf life calculation according to the definition given in ICH Q1E.

The storage condition involves both temperature and humidity levels. For drug product, the stability protocol calls for collection of samples held at the given storage condition at fixed time intervals for chemical and physical analysis. A typical stability study will have 3 batches of final drug product studied across several factor combinations according to a factorial design. When the design is fractional, by convention the design is referred to as a "matrix" design, or if the extremes of the range are chosen, a "bracket" design. Statistical modeling is carried out according to a batch-specific linear model, with rules given for pooling parameter estimates across the batches. In the following sections, we will review the ICH Q1E batch specific model, the rules for pooling, a critique of the rules from an empirical and scientific perspective and end with a recommendation. We will focus specifically on the case of the active assay for small molecule drug products.

2. ICH Q1E Model and Pooling Rules

2.1 Batch-Specific Model and Shelf life

The following fixed effects linear model is used to describe the stability profile for each batch, assuming a batch-specific initial and rate of change and common error across the batches:

$$y_{ij} = A_i + B_i \times T_{ij} + \varepsilon_{ij}, \quad \text{Model 1}$$

where y_{ij} = API concentration (in %label) for the *i-th* batch and *j-th* time point,
A_i = intercept corresponding to *i-th* batch at time 0,
B_i = rate of change for the *i-th* batch,
T_{ij} = *j-th* time point for *i-th* batch,
ε_{ij} = residual error, $\varepsilon_{ij} \sim N(0, \sigma_e^2)$.

The shelf life is calculated based on the intersection of the lower specification limit (*LSL*) and one-sided lower 95% [$=100\times(1-\alpha)\%$] confidence limit. Thus the shelf life, S_i, of the i[th] batch is the solution to T_{ij} in the following equation:

$$LSL = A_i + B_i \times T_{ij} - t_{\alpha,df} \times \sqrt{Var(A_i + B_i \times T_{ij})}$$, if $B_i < 0$,

where A_i and B_i are as defined previously, Var(A_i+B_i×T_{ij}) is the variance of the estimate at time T_{ij} (from Model 1), and $t_{\alpha,df}$ is the appropriate Student's t-quantile satisfying $P\{U_{df} > t_{\alpha,df}\} = \alpha(=.05)$ where U_{df} follows Student's t-distribution with degrees of freedom = *df*.

2.2 Pooling Rules according to ICH Q1E

The ICH Q1E pooling rules proceed in steps as follows with each test carried out against p-value=0.25.

1. Test slopes
   - 1.1 If slopes pass
     - 1.1.1 Test intercepts
       - 1.1.1.1 If intercepts pass, fit common model across all batches, single shelf life calculated
       - 1.1.1.2 If intercepts fail, fit common slope, batch-specific intercepts model, calculate shelf life for each batch, report most conservative shelf life.
   - 1.2 If slopes fail
     - 1.2.1 Test intercepts
       - 1.2.1.1 If intercepts pass, fit batch-specific slope, common intercept model, calculate shelf life for each batch, report most conservative shelf life
       - 1.2.1.2 If intercepts fail, fit batch-specific slope and batch-specific intercept models, calculate shelf life for each batch, report most conservative shelf life.

2.3 Example of NDA Stability Study

To place the ICH Q1E pooling rules in better perspective, it may be instructive to consider the design of a not atypical stability protocol the authors reviewed recently. The protocol described the analytical requirements for an immediate release tablet drug product manufactured at strengths 25 mg, 75 mg, 100 mg, 125 mg and 150 mg from a direct compression (DC) process and packaged in five configurations of HDPE Bottles and Aclar Blister. Three distinct drug substance batches would be used to manufacture 9 batches of final product tablets representing the 5 strengths according to the following bracketed matrix design given in Table 1.

Table 1 – Correspondence of Drug Product Batches to Drug Substance Batches

Drug Substance Batch	Strength (mg)				
	25	75	100	125	150
DS 1	X	X			X
DS 2	X		X		X
DS 3	X			X	X

Each "X" in Table 1 represents a batch of drug product. Each of the 9 batches would be studied at 6 different temperature-humidity storage conditions and packaged in 5 different configurations. Each batch would generate 30 different stability profiles. A stability model of the resulting data capturing the initial concentrations and rates of change would contain 279 parameters. Can we reasonably apply the ICH Q1E rules to such a large number of parameters, especially given that independence traces back ultimately to only 3 drug substance batches?

3. Critique of the ICH Q1E Pooling Rules

In the previous section 2.3 where an example NDA stability protocol design was discussed, the operational challenges of applying the rules to a set of data that contains 279 parameters was alluded to, keeping in mind that independence between parameters was really only at the drug substance level with only 2 degrees of freedom. The 9 final drug product batches cannot be modeled assuming independence between their parameters given the manufacturing design.

We emphasize that the basic objective of an NDA stability study is to characterize 'chemical stability' or a degradation rate (a rate of change). In a sense, we can regard it as a chemical kinetic study with respect to the API in the presence of known fixed effects specified in the protocol, such as storage condition and package. Emphasis is on the stability behavior at the batch mean level. This is particularly evident in the use of composite samples and is completely consistent with the clinical development paradigm where dosage units cannot be tested individually prior to being administered to subjects. Batch characterization in terms of the mean of the batch must be the standard quality statement, with additional compendial or other Content Uniformity testing to assure individual dosage unit delivery of potency.

It should be noted that statistical input on the powering and design of such studies with respect to fixed or random factors is not standard practice. The emphasis is on

economical estimation of the effects of multiple storage conditions, packages, and strengths as fixed effects through a small number of batches. The current designs carried out for this purpose have stood the test of time and are adequate for the purpose of parameter estimation. Another important concern is the design of the chemical analytical methodology. This is rarely considered in the stability protocol design, yet it has the potential to confound fixed effects with analytical runs. This adds additional lack of independence to the stability profile estimation. It is desirable to consider between and within analytical run variance components in properly assessing the uncertainty in the estimation of the stability model parameters.

The test for common intercepts is particularly problematic. It is totally unrealistic to assume batch potencies are identical at release. There will always be small differences in physical weighing operations and blending which will cause batches to be different. That is a reasonable *a priori* expectation, so it flies in the face of the Q1E rules which seek to show 'equality'. How does one justify the use of a residual error term that is essentially derived from measurement error to be the standard for poolability across batch intercepts? Batch dispersion at time of manufacture is expected to be exacerbated by analytical measurement variability. In addition, the hypothesis testing approach is a disincentive to the pursuit of precise measurement tools and analytical methods.

The test for poolability of slopes across batches is also problematic. Essentially, what it admits is that the chemistry and physical processes governing the batch production process is dependent on the batch. This is again a harsh view of modern manufacturing technology where control of process parameters and materials is maintained at a very high level. Given these circumstances, it follows that the chemistry should be generally independent of batch, so the assumption of a common fixed rate constant is reasonable.

We also offer empirical evidence in support of the common slopes model as a reasonable assumption in practice. We reviewed 33 recent stability studies representing a broad range of development and marketed compounds in various dosage forms. We fit a fixed by batch model and tested for poolability of the slopes assuming batch-specific intercept. We found a median p-value for the test of poolability = 0.670, with 13% of p-values <0.25, 5% of p-values <0.10 and 70% of p-values >0.50. This distribution of p-values is probably not much different from what one might expect if the test results were completely due to chance alone approximating a uniform distribution. There was no evidence in this empirical study to suggest that there is a general tendency for the existence of batch-specific slopes.

Some have advocated an equivalence approach to pooling (Tsong et al, 2003), This merely moves the question of pooling from one hypothesis test procedure to another. We believe that hypothesis testing in the context of stability modeling is misplaced given that the objective is estimation of a rate of change. The inability to power stability studies further calls into question the wisdom of applying an equivalence approach. Even more, what would an appropriate equivalence criterion be?

4. Mixed Model and Bayesian Approaches

4.1 Basic Model

The main objective of a stability study from the statistical perspective is to estimate the

parameters of the statistical model, account for incipient variation and acknowledge the independence structure in a way that control over the batch mean is assured. We submit that a reasonable way to achieve this statistical goal is through the use of standard mixed effects modeling (Fiztmaurice et al, 2004; Littell et al, 2006). The mixed model is consistent with the basic philosophy that batches arise from a fixed manufacturing process. Batches can be regarded as the primary independent statistical units as subject effects. The mixed modeling also acknowledges all sources of variation through a simple but flexible variance structure that is easily interpretable. It is a natural representation of a batch manufacturing process and directly leads to process simulations and post commercialization studies to propose and confirm control strategies. The model is also extendable to multiple fixed factors under study as well as multivariate responses.

Table 2 provides two mixed models which are useful for evaluating stability study data for the purpose of characterizing the stability profiles where the subscripts *i, j, k* correspond to Batch, Storage Condition and Time respectively. The model is easily extendable to include the effects of fixed factors such as package or manufacturing site. In practice, for small molecule compounds, Model A is sufficient for shelf life calculation, control limits calculations and simulations.

Table 2 – Two Forms of Mixed Models

Model	Form	Number of Parameters	
		Fixed Effects	Variance
A	$y_{ijk} = (\mathrm{A}_0 + \alpha_i) + B_j \times T_{ijk} + \varepsilon_{ijk}$	n_c+1	2
B	$y_{ijk} = (\mathrm{A}_0 + \alpha_i) + (B_j + \beta_i) \times T_{ijk} + \varepsilon_{ijk}$	n_c+1	3/4[a]

Index: *i*=Batch, *j*=Condition (j=1,2,..n_c), *k*=Time.
[a] if correlated random terms in intercept and slope

One objection to the use of the mixed model is the concern that during development, only a small number of batches are typically available. Although we understand why this is a concern, having only a small number of batches in no way diminishes the practical value of the mixed model in assessing stability profiles. This is for the reasons discussed earlier, with the most important reason being that it captures the important parameters of a batch manufacturing process in a concise and easily interpreted manner. Furthermore, as discussed in the next section, a Bayesian approach incorporating knowledge and experience gained from similar products and processes from process experts can to some extent mitigate the concern with small numbers of batches and improve the modeling.

4.2 Bayesian Approach

The Bayesian approach (Carlin et al, 2009) provides a mechanism to include prior information to the statistical analysis of data and to update model parameter estimates as new data are collected. In addition, a Bayesian posterior predictive calculation leads to a more natural way to enable strategies in relation to controlling risk of Out of specifications. Risk calculations would have a direct interpretation as a probability in this context. Currently available software such as WinBUGS and SAS 9.3 provide readily available tools to carry out Bayesian calculations.

The Bayesian approach requires specification of a prior distribution on the unknown parameters, μ corresponding to the process average at time 0; β_j corresponding to the rate

of change at j^{th} storage condition; σ_{ε}^2 corresponding to the residual error variance; σ_{α}^2 corresponding to manufacturing variability or batch to batch variability. As an example of integrating expert opinion, we used the following information from the process engineers:

- Process mean is likely between 99% and 101%, translated into the statistical statement that $\mu \sim N(100, 0.1)$,
- Manufacturing or batch variance is likely between 0.1 and 0.5, translated into the statistical statement $\sigma_{\alpha}^2 \sim \Gamma^{-1}(10, 2)$
- No prior information was available on the rate parameter so a uniform prior was chosen for the rates of change, translated into a statistical statement as $\rho(\beta_j) \propto 1$,
- Residual variance consisting primarily of analytical method uncertainty involving repeatability and intermediate precisions estimates is likely between 0.1 to 1.0, translated into a statistical statement as $\sigma_{\varepsilon}^2 \sim \Gamma^{-1}(6, 2)$.

4.3 Stability Times as Random Blocks

Another advantage of the mixed model is that it is possible to directly model the block structure inherent in stability pulls. Frequently in stability studies, we find that resource limitations will cause blocking across multiple batches to be analyzed in one or a few number of analytical runs. This induces a dependency in the stability profiles across the different batches. The mixed model allows this dependency within analytical runs at specific time points to be estimated by calculating between and within analytical runs variance components, where different time points are understood to constitute the between runs component. One of the consequences of such a model is that the uncertainty measurement associated with the fixed parameter estimates will have a more complicated form. A Satterthwaite or Kenward-Roger adjustment on the degrees of freedom associated with the uncertainty is an acceptable approach. However, the Bayesian approach would eliminate this need for an approximation, so from that perspective, the Bayesian approach is also preferable.

5. Recommendation and Conclusions

Current regulatory guidelines for assessing stability and shelf life claims are being challenged in view of current technologies and scientific understanding. The poolability tests described in the ICH Q1E guidance serve statistical convenience needs and are not rooted in a firm understanding of batch manufacturing considerations and science. As the pharmaceutical industry moves more deeply into a Quality by Design paradigm for pharmaceutical manufacturing, scientific and engineering principles will increasingly drive the development and commercialization of products. It is in this light that we believe this is the right time to challenge the pooling rules embodied in ICH Q1E. Should they be reconsidered in light of scientific and engineering considerations as well as current computing technologies? Based on empirical and scientific considerations, we believe that a common fixed slope for assay independent of batch is not an unreasonable assumption in many stability studies given the conditions outlined previously. We recommend a mixed effects model as a more natural representation of a fixed manufacturing process compared to the ICH Q1E fixed effects model. A Bayesian framework can incorporate process engineering and scientific judgment.

References

Carlin,B.P., Louis,T.A. (2009). *Bayesian Methods for Data Analysis.* Boca Raton FL:CRC Press.

International Conference on Harmonization (2003*).* EVALUATION FOR STABILITY DATA Q1E. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q1E/Step4/Q1E_Guideline.pdf

Littell,R.C., Milliken,G.A., Stroup,W.W., Wolfinger,R.D., Schabenberger,O. (2006). *SAS for Mixed Models* (Second Edition). Cary,NC: SAS Institue

Lunn, D.J., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing, 10:325-337.

SAS Institute Inc. (2011) SAS/STAT® 9.3 *User's Guide* Cary, NC: SAS Institute Inc.

Tsong,Y., Chen, W-J., Lin,T-Y.D., Chen,C.W. (2003). Shelf Life determination Based on Equivalence Assessment. J. of Biopharm.Stat., 13(3):441-449.