Current Practices in Shelf Life Estimation

James Schwenke
Boehringer Ingelheim Pharmaceuticals, Inc.
PQRI Stability Shelf Life Working Group

Shelf Life Estimation

- ICH estimation methods
- Typical stability study
- Shelf Life paradigm
- Regression methods for estimating shelf life
- Actual stability study results
- Summarize empirical distribution of estimated shelf life
- Extend to random batch analyses

Definitions of Shelf Life – ICH Q1E

- ICH Guideline Q1E defines shelf life as
  "The shelf life of a pharmaceutical product is the maximum time at which the true mean response of a stability limiting characteristic crosses the acceptance criterion."

  - basis for the current ICH/FDA shelf life estimation procedure
  - limited assurance that individual test results will comply with the specification up to m months
  - focus on the mean response implies the risk to fail specification at shelf life will be 50%

Typical Stability Study

- Minimum of 3 stability batches
  - can be 6 or more batches included in study
  - can be several studies combined together
- Study duration can be 6-12-24-36-48 months
  - longer studies can have interim reports
  - length of study can depend on shelf life desired
- Various environmental conditions
  - more severe environmental condition can act as an accelerated testing for milder conditions

The Shelf Life Paradigm

- Response
- Acceptance Criterion
- Storage Time (Months)
  - Product Shelf Life
- Stability Limiting Characteristic
- Acceptance Criterion

- The shelf life of a pharmaceutical product is the maximum time at which the true mean response of a stability limiting characteristic crosses the acceptance criterion.

  - basis for the current ICH/FDA shelf life estimation procedure
  - limited assurance that individual test results will comply with the specification up to m months
  - focus on the mean response implies the risk to fail specification at shelf life will be 50%
The Shelf Life Paradigm

- General concerns
  - Exactly what should be modeled and how does it relate to the product shelf life?
    - individual tablet
    - composite sample of several tablets
    - packaged unit (bottle or blister pack)
    - stability batches or all future batches
  - How does results of content uniformity studies affect the decision process?
    - If focused on the mean response, when the mean crosses the acceptance criteria, 50% of product out of specification.

ICH Shelf Life Estimation Methods

- ICH methodology
  - minimum of three batches
  - batches are considered fixed effects for the analysis
  - batches can be pooled if no significant differences
    - 0.25 level of significance for tests involving batches
    - 0.05 level of significance for tests involving other factors
      - package type, storage orientation, coating, etc.
  - construct 95% confidence intervals on individual (or pooled) batch means
  - find minimal storage time where confidence interval crosses acceptance criteria

ICH Estimation Methodology

- usually a simple linear (straight line) regression model is assumed to characterize the response-time continuum
  - first-order nonlinear models are more appropriate for some stability limiting characteristics
  - we will focus on simple linear regression models

  simple linear model
  \[ y_{ij} = b_{0i} + b_{1i} (\text{month}_j) + \epsilon_{ij} \]
  - \( y_{ij} \) = observation at \( j \)th month for \( i \)th batch
  - \( b_{0i} \) = batch intercept and \( b_{1i} \) = batch regression slope
  - \( \epsilon_{ij} \) = residual error with Normal assumptions

- ICH methodology suggest to test for equal regression slopes among stability batches first
  - if batch slopes are nonsignificant (\( \alpha = 0.25 \))
    - common regression slope is assumed among batches
    - batch intercepts are tested
      - if batch intercepts are nonsignificant (\( \alpha = 0.25 \))
        - common intercept is assumed among batches
        - batches are pooled
  - if batch slopes are significant, no further testing is considered
  - ICH does not allow for a model with a common batch intercept and unequal slopes (which can be an important model)
ICH Estimation Issues

• Regression model selection
  - there are four possible linear regression models
    1) full model: unequal intercepts and slopes among batches
    2) common intercept with unequal slopes among batches
    3) unequal intercepts with common slope among batches
    4) common intercept and slope (pooled batches)
  - Model #2 is not allowed following ICH guidelines
    ◦ still considered an important model to consider by colleagues
    ◦ for stability limiting characteristics that should be at 0% or 100% at 0-months storage time

• Batch poolability
  - If batches cannot be pooled, shelf life is estimated on results of the worst batch
  - If batches can be pooled, between and within batch variation is combined

• Random batch effects
  - the 0.25 level of significance used to test hypotheses involving fixed batch effects is intended to accommodate batch-to-batch variation
  - available software allows for random batch analysis
  - would avoid batch poolability issue

Real-Life Example Data Set

• Real-life example contributed by one of our PQRI members
  - 26 stability batches
  - all on same product
  - most kept on study for 24 months
  - assay was measured
  - will use to study empirical distributional properties of estimated shelf life using 3- and 6-batch studies

Product Shelf Life

Empirical Study of Distributional Properties

• use real-life stability batch data set
  - consider entire 26-batch data set defines the product population of batches
  - use all batches to product shelf life
    ◦ estimate regression line assuming batches are random
    ◦ product shelf life is storage time where regression line crosses acceptance criteria
  - consider all possible sets of 3 and 6 batches
    - conduct ICH estimation methods for shelf life
    - summarize results

Empirical Study of Distributional Properties

• 3-batch analysis
  - consider all possible combinations of 3 batches from the 26
    - there are 2,600 combinations
    - conduct regression analysis allowing for all four models
      ◦ does not follow ICH
      ◦ allows for common intercept / unequal slope model
      ◦ model is included by my analytical scientists
      - estimate shelf life from best fitted model
  - 6-batch analysis
    - there are 230,230 possible combinations (7.5 days to run)
    - randomly chose 20,000 (15.5 hours to run)
ICH Estimation Methodology

- Comparing the two empirical distributions
  - there is a shifting toward shorter estimated shelf lives with an increase in the number of batches included in the analysis
    - counterintuitive
    - increase in the number of batches should reflect an increase in the amount of information about the product
    - increase in the amount of information about the product should reflect a better estimate of shelf life
  - should see a shift in distribution toward longer shelf lives
  - better estimates of product shelf life (37.8 months)
  - disincentive for industry to include more stability batches

3-Batch Estimate of Shelf Life

- n = 466 (18%) mean = 22.9 months SD = 5.86

6-Batch Estimate of Shelf Life

- n = 5035 (25%) mean = 19.9 months SD = 4.60

3-Batch Estimate of Shelf Life

- n = 788 (30%) mean = 23.3 months SD = 5.71
**ICH Estimation Methodology**

- **Poolability**
  - Concept is to estimate shelf life on the best fitted regression
  - Accommodate random variation among batches by allowing pooling of batch data through regression parameter estimates
  - Allow for common slope and intercept models to characterize batch response
  - Use $\alpha = 0.25$ level of significance
  - Estimate of shelf life too heavily dependent on best model
  - Assuming unequal slopes forces the shelf life estimate to be based on worst batch
  - Tends to minimize shelf life estimate
Random Batch Mixed Model Analysis

- Two rationale to suggest an alternative random batch analysis
  - can extend inference of estimated shelf life to future batches
  - avoids dependence of shelf life estimate on “best” model fit
- mixed model analysis would
  - model between-batch variation as a random effect
  - quantify both between and within-batch variation separately
  - allows broad and narrow inferences
  - allows estimation of shelf life through calibration techniques
    - defined by a one-sided (lower) interval estimate on calibration storage time point

Reflection Calibration Method (Total = 2600)

3-Batch Estimate of Shelf Life – Confidence Interval

- Using 3 Batches with 24 Months of Data, Extrapolating to 48 Months of Storage Time
- Estimating Shelf Life Using Random Batch Mixed Model Analysis
- Product Shelf Life = 37.8
- 12.46%
- 13.69%
- 14.46%

6-Batch Estimate of Shelf Life – Confidence Interval

- Using 6 Batches with 24 Months of Data, Extrapolating to 48 Months of Storage Time
- Estimating Shelf Life Using Random Batch Mixed Model Analysis
- Product Shelf Life = 37.8
- 13.59%
- 11.37%
- 15.76%
- 16.47%
- 15.01%

Random Batch Mixed Models – Distribution of $\hat{\beta}$

- Acceptance Criteria
- Mean of Random Batch Mixed Model Analysis
- Labeled Shelf Life

6-Batch Estimate of Shelf Life – Confidence Interval

- Using 6 Batches with 24 Months of Data, Extrapolating to 48 Months of Storage Time
- Estimating Shelf Life Using Random Batch Mixed Model Analysis
- Product Shelf Life = 37.8
- 13.59%
- 11.37%
- 15.76%
- 16.47%
- 15.01%

Random Batch Mixed Models – Reflection

- Acceptance Criteria
- Confidence Band
- Mean of Random Batch Mixed Model Analysis
- Labeled Shelf Life
6-Batch Estimate of Shelf Life – Confidence Interval

Alternatives to ICH Shelf Life Estimation Methodology Using Industry Data
Using 6 Batches with 24 Months of Data, Extrapolating to 48 Months of Storage Time
Estimating Shelf Life Using Random Batch Mixed Model Analysis

Product Shelf Life = 37.8

Conclusions

• Moving from a fixed-batch to a random-batch analysis has several advantages
  – can be used to extend the inference of a shelf life statement to future batches
    - depending on how well the stability data represent the manufacturing process
  – avoids the batch poolability issue
    - breaks the dependence of the estimate of shelf life on which reduced or pooled regression model is selected
    - avoids the issue of a common intercept / different slope regression model being appropriate or not

• Better quantifies between and within-batch variation by using an appropriate statistical model
• Allows additional stability batches to benefit the estimate of shelf life
    - for fixed-batch analysis, additional stability batches increases chance that shelf life is estimated by worst-batch
    - for random-batch analysis, additional stability batches adds to the information on between-batch variation
• Better distributional characteristics of shelf life estimate

Conclusions

• Random-batch analysis
  – as presented, is still estimating an “average” batch shelf life
    - confidence interval approach was discussed which is consistent with ICH
    - still have conceptual problem that at shelf life, half of product unit is above specification limit
    - does not reflect desired quality statement
  – alternatives methods for shelf life estimation
    - tolerance intervals
    - quantile regression