# Shelf Life Estimation: Bayesian Augmented Mixed Model Approach

Maryna Ptukhina[1], Walter Stroup[1]
[1] Department of Statistics University of Nebraska -Lincoln, 340 Hardin Hall North Wing Lincoln, NE 68583-0963

## Abstract

Shelf life estimation procedures, following ICH guidelines, use multiple batch regression with fixed batch effects. This guidance specifically mandates estimates based on at least 3 batches. Technically, the fixed-batch model limits inference to the batches actually observed, whereas ICH requires resulting estimates to apply to all future batches stored under similar conditions. This creates a conflict between the model used and the inference space the model is intended to address. Quinlan, et al. (2013) and Schwenke (2010) studied the small sample behavior of this procedure. Both studies revealed large sampling variation associated with the ICH procedure, producing a substantial proportion of extremely low and extremely high estimates. Quinlan, et. al (2013) also considered alternative approaches including mixed models with random batch effects. While this eliminated the conflict between model and intended inference space, there were still problems with the mixed model approaches Quinlan considered. We present a Bayesian augmented mixed model approach to shelf life estimation that takes advantage of the theoretical benefits of the mixed model and uses prior information about variance components to improve accuracy of shelf life estimation procedure.

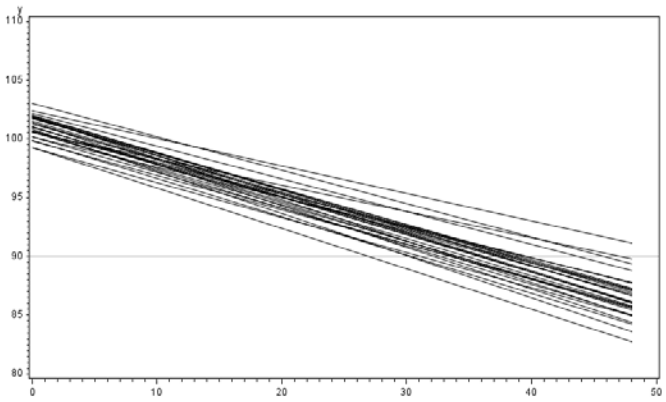
**Key Words:** shelf life estimation, BLUP

## 1. Introduction

Accurate shelf life estimation is very important to a variety of applications. This paper specifically focuses on the pharmaceutical industry, where inaccurate estimation can lead to undesirable consequences. Overestimation of shelf life could lead to consumption of drugs that are no longer stable and effective, while underestimation can cause the consumer to discard good product prematurely. Thus accurate estimation of shelf life is essential to both consumers and producers. Recent research suggests that shelf life estimation procedures used in the pharmaceutical industry are not always reliable: overestimation and underestimation are common. This paper explores new techniques that utilize prior information gained through previous stages during the development process, thus minimizing the required number of replications. These techniques are geared to provide consumers with accurate shelf life estimates while maintaining lower costs for the production facilities**.**

## 2. Shelf life

Suppose we have a product that deteriorates over time. In a pharmaceutical setting, this could be a drug or a vaccine. In the simplest case, we could imagine that product's effectiveness decreases linearly over time, and its lifetime is determined by how long its measure of effectiveness remains within a defined acceptance criterion. Figure 1 shows a population of batches, where y is the mean response of a stability limiting characteristic and x is the storage time in months.

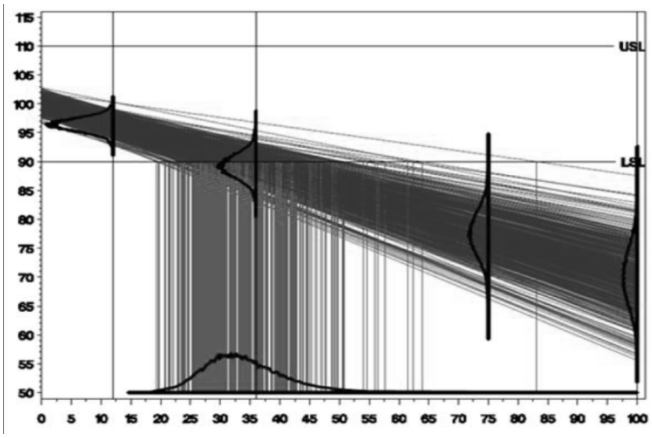



Acceptance criterion=90

**Figure 1:** Distribution of batch mean responses over time

It is important to note that when we have a distribution of mean batch responses not all intercepts are the same and not all slopes are the same. Instead it is reasonable to assume that there is random variation among intercepts and slopes.

Once the distribution of y is specified, the distribution of shelf lives arises as a consequence. When a batch hits the acceptance criterion, in this case A=90, the value gets projected on the horizontal axis. When we do this for each batch we get a distribution of shelf life on the horizontal axis.
Figure 2 illustrates this relationship between the batch response distribution and shelf life distribution.




Acceptance criterion=90

**Figure 2[1]:** Distribution of batch mean responses over time and distribution of shelf lives

[1] Picture credited to Quinlan (2010).

In statistics we often focus on the mean and median of the distribution. However if we focus on the median, for example, the statement that we would be able to make is that with 50% probability any given batch's shelf life will meet or exceed the shelf life established for the product. We would like that probability to be at least 95%. Therefore, instead of the median we will focus on 0-5$^{th}$ percentile of the distribution of shelf life**.**

## 2.1 Fixed batch effects model

A model typically used for this type of problem can be written as

$$y_{ijk} = \beta_0 + b_{0i} + (\beta_1 + b_{1i}) X_j + e_{ijk}.$$

This paper will focus on the implications and resulting behavior of different assumptions about $b_{0i}$ and $b_{1i}$ as well as different ways of implementing these assumptions.

Pharmaceutical shelf life is currently estimated following ICH guidelines that define batch effects (i.e. $b_{0i}$ and $b_{1i}$ ) as fixed.

Quinlan, et al. (2013) and Schwenke (2010) studied the small sample behavior of this procedure. Both studies revealed large sampling variation associated with the ICH procedure, producing a substantial proportion of extremely low and extremely high estimates.

## 2.2 Mixed model

ICH also requires resulting estimates to apply to all future batches stored under similar conditions. This creates a conflict between the fixed batch effects model used and the inference space the model is intended to address.

A random coefficient linear mixed model provides an alternative to the current approach, since it allows for inference to be applied to all future batches.

Random coefficient linear mixed model is given by $y_{ijk} = \beta_0 + b_{0i} + (\beta_1 + b_{1i}) X_j + e_{ijk}$ , where $b_{0i}$ and $b_{1i}$ are assumed to be random with the following distribution

$$\begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix} \right).$$

Often in practice we assume $\sigma_{01}$ is equal to zero.

Quinlan, et. al (2013) considered this mixed model with random batch effects. While this eliminated the conflict between model and intended inference space, there were still problems with the mixed model approaches that were considered. Specifically, while the variability of the mixed model approach was improved compared to ICH procedure, the mixed model approach produced a relatively large proportion of estimates that were too high.

Stroup and Quinlan (2015, in press) have explored this mixed model methodology further and found that shelf life estimates based on the lower BLUPs of batch specific slope and intercept of the shortest-lived batch were performing the best out of several other competing mixed model based approaches.

The question motivating this study is: "What if we try a Bayesian approach to shelf life estimation"? In other words would a use of plausible, somewhat informative priors improve the estimation procedure? If we have information about variance components from the previous stages of drug development we could use it to inform the priors. To answer these questions we performed a simulation study that is described in the next section. Before we describe the simulation and its results let us look at different ways of calculating shelf life.

## 2.3 How can we calculate shelf life?

Using the random coefficient linear regression mixed model given above we could get the estimates for $\hat{\beta}_0$ and $\hat{\beta}_1$ using the mixed models estimation procedure.

Estimated shelf life is then $\hat{T} = \dfrac{A - \hat{\beta}_0}{\hat{\beta}_1}$, where A is the acceptance criterion.

Quinlan has tried this approach and showed that the resulting shelf life estimates were too high.

Instead of using the estimates for $\hat{\beta}_0$ and $\hat{\beta}_1$, we could get the confidence bounds for $\hat{\beta}_0$ and $\hat{\beta}_1$ using the mixed models estimation procedure and calculate shelf life based on the lower confidence bounds. Estimated shelf life is then $\hat{T} = \dfrac{A - \hat{\beta}_{0,L}}{\hat{\beta}_{1,L}}$,

where A is the acceptance criterion, $\hat{\beta}_{0,\mathrm{L}}$ and $\hat{\beta}_{1,\mathrm{L}}$ are the lower bounds for slope and intercept. For example, we can use the 5th percentile of $\hat{\beta}_{0,\mathrm{L}}$ and $\hat{\beta}_{1,\mathrm{L}}$, which are the lower bounds of a one sided 95% confidence interval.

Also, we could get the shelf life estimate on the BLUP using the batch-specific regression equation $\hat{\beta}_0 + \hat{b}_{0i} + \left(\hat{\beta}_1 + \hat{b}_{1i}\right) X$. To do that we would identify the batch-specific regression with the shortest shelf life and use it as a basis for the shelf life estimate.

Estimated shelf life is then $\hat{T} = \dfrac{A - \left(\hat{\beta}_0 + \hat{b}_{0i}\right)}{\left(\hat{\beta}_1 + \hat{b}_{1i}\right)}$.

Or, alternatively, we could base the shelf life estimate on the lower confidence bounds of the batch-specific intercept and slope BLUPs of the batch with the shortest shelf life.
This is similar to the method described above, but instead of using $\hat{\beta}_{0,\mathrm{L}}$ and $\hat{\beta}_{1,\mathrm{L}}$ we use $\left(\hat{\beta}_0 + \hat{b}_{0i}\right)_L$ and $\left(\hat{\beta}_1 + \hat{b}_{1i}\right)_L$.

# 3. Simulation

## 3.1 Empirical shelf life distribution

The first step was to generate the empirical shelf life distribution. One thousand data sets were generated using a random coefficient regression model with 3 batches per trial. Observations were generated for each batch at times 0, 3, 6, 9, 12, 18 and 24 months.
Parameters used in simulation were similar to Quinlan's $\beta_0 + b_{0i} \sim NI(101, 1.5)$ and $\beta_1 + b_{1i} \sim NI(-0.33, 0.0015)$, where $b_{0i}$ and $b_{1i}$ are random intercept and slope effects,

assumed to be uncorrelated. The within-batch random variability is $u_{ij}$, which has distribution: $u_{ij} \sim NI(0, 0.5)$. The acceptance criterion was set to 90.
A histogram of the empirical distribution of shelf lives is shown on Figure 3 below.

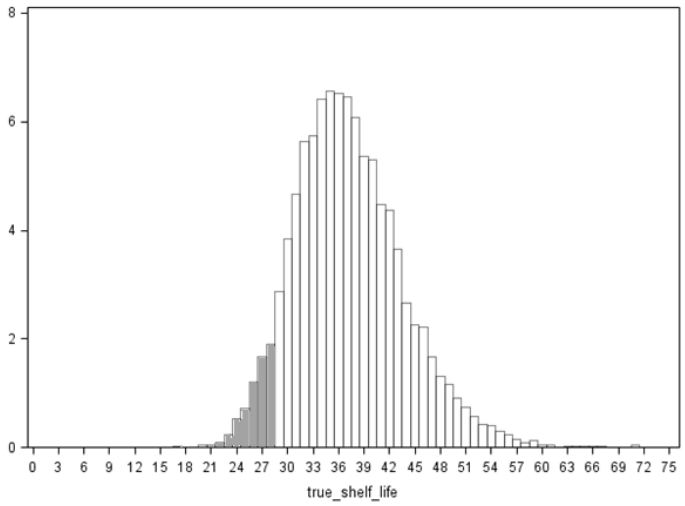



**Figure 3:** Empirical distribution of shelf lives

From Figure 3 we can see that the empirical distribution of the "true" shelf life ranges from 17 to 75 months with median at 37 months. From discussion in section 2 recall that we are interested in values from 0 to 5$^{th}$ percentile, since the statement that we would like to make is that at least 95% of all future batches will meet the acceptance criterion.

For the distribution shown in Figure 3 this implies a range of 17 to 28 months. Thus, to evaluate the performance of different estimating procedures we would consider the range of values between 17 and 28 months as desirable estimates of shelf life.

## 3.2 Estimation procedures

As mentioned earlier the idea behind a Bayesian augmented mixed model approach is to use the information that we have from the previous stages of product development and try different combination of priors to learn about the behavior of shelf life estimates for each scenario.

To explore this we decided to look at four different scenarios shown in the Table 1 below. Technically, priors on $\sigma^2_0$ and $\sigma^2_1$ should be called hyperpriors, but for simplicity of this discussion they will be referred to as priors.

**Table1:** Four cases of priors

| | Vague priors on $b_{0i}, b_{1i}$ | Informative priors on $b_{0i}, b_{1i}$ |
|---|---|---|
| **Vague priors on** $\sigma_0^2, \sigma_1^2, \sigma^2$ | **Case I** | **Case IV** |
| **Informative priors on** $\sigma_0^2, \sigma_1^2, \sigma^2$ | **Case II** | **Case III** |

For each scenario we tried different sets of vague and informative priors, the results presented for each case are typical results that are representative of each set of prior combinations. Results for each scenario are described below. In this section lower confidence bounds of 95% confidence interval of mixed model estimates $\hat{\beta}_{0,L}$ and $\hat{\beta}_{1,L}$ were used to get estimates of shelf life.

Simulations were performed is SAS 9.4 using PROC MCMC™ . The number of Markov chains used for each scenario was 2,000,000 with a burn-in of 100,000 and thinning of 1000.

### *3.2.1 Case I: Vague priors*

A typical distribution of shelf life estimates produced using vague priors on both sets of parameters is shown in Figure 4. We can note that resulting shelf life estimates are lower than what would be considered a desirable result, i.e. lower than 17 months. For this specific result the following set of vague priors was used: improper uniform-like flat prior for the slope and intercept (in SAS it is referred to as general (0) distribution) and inverse gamma distribution for the variance components $\sigma^2 \sim IG(0.01, 0.01)$.

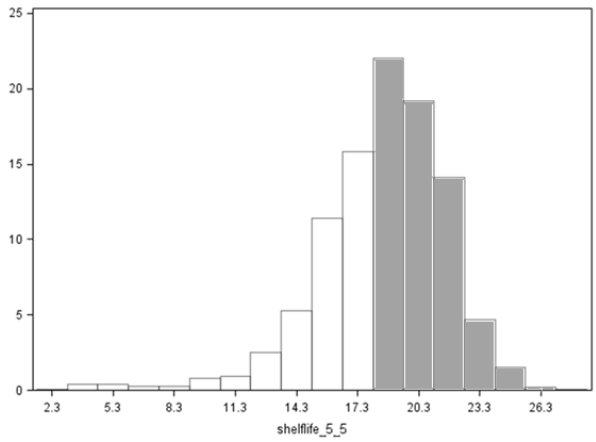



**Figure 4:** The resulting distribution of shelf life estimates using vague priors

From Figure 4 we can see that approximately 40% of estimates of shelf life are much lower than the desired shaded range between 17 and 28 months.

### *3.2.2 Case II: Vague priors for b, informative for $\sigma^2$*

Using vague priors for the slope and intercept and informative priors for the variance components typically resulted in shelf life estimates which were higher compared to the Case I scenario, but this method produced a relatively large proportion of estimates that were too high, i.e. greater than 28 months.

Figure 5 shows the results based on a general (0) prior for b and three different inverse gamma priors for each variance component : $\sigma^2{}_0 \sim IG(3,6)$ , $\sigma^2{}_1 \sim IG(3, 0.006)$, $\sigma^2 \sim IG(3,2)$. Parameter values for each inverse gamma distribution were chosen by supposing that information in previous stages of development suggests that the most likely values of variance components were $\sigma^2{}_0$=1.5, $\sigma^2{}_1$ = 0.0015 and $\sigma^2$=0.5 respectively.

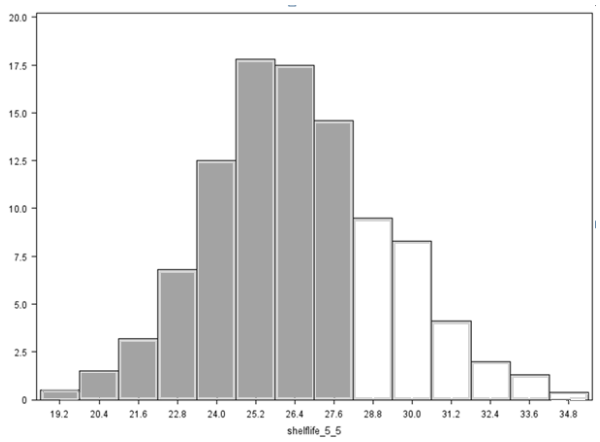



**Figure 5:** The resulting distribution of shelf life estimates using vague priors for b, but informative priors for $\sigma^2$

The desirable range of estimates between 17 and 28 months is shaded in grey. In this case 25% of the estimates of shelf life were too high.

### *3.2.3 Case III: Informative priors*

Using informative priors for the slope, intercept and variance components typically led to shelf life estimates which were higher compared to the Case I scenario, yet similar to the Case II scenario. This method produced a relatively large proportion of etimates that were too high.

Figure 6 shows the results based on the following priors: $b_0 \sim N(100, 9), b_1 \sim N(-0.33, 0.05)$ and three different inverse gamma priors for each variance component: $\sigma^2{}_0 \sim IG(5,9)$, $\sigma^2{}_1 \sim IG(5, 0.009)$, $\sigma^2 \sim IG(5,3)$.

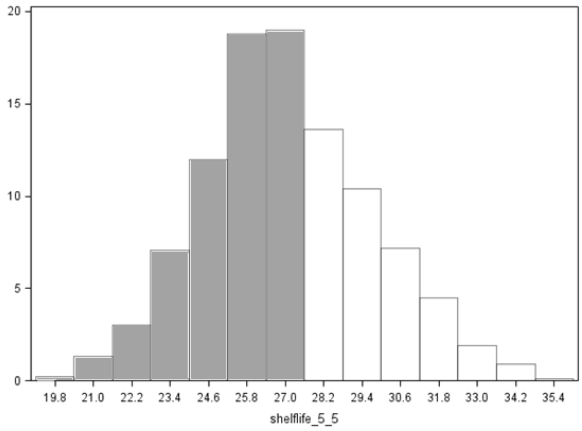



**Figure 6:** The resulting distribution of shelf life estimates using informative priors

The desirable range of estimates between 17 and 28 months is shaded in grey. In this case about 45% of the shelf life estimates were too high.

### *3.2.4 Case IV: Informative priors for b, vague priors for* $\sigma^2$

Typical distribution of shelf life estimates produced using informative priors for slope and intercept, but vague priors for variances is shown in Figure 7. We can note that resulting shelf life estimates are very close to a desirable result.
Figure 7 shows results using the following priors: $b_0 \sim N(100, 9), b_1 \sim N(-0.33, 0.05)$ and common prior for $\sigma^2 \sim IG(0.01, 0.01)$. The area shaded in grey represents the desirable shelf life estimates between 17 and 28 months and covers approximately 98% of distribution of shelf life estimates.

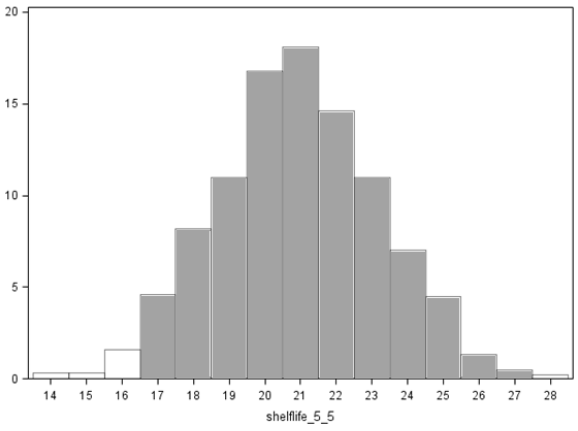



**Figure 7:** The resulting distribution of shelf life estimates using informative priors for b, but vague priors for $\sigma^2$

Since the results of the fourth scenario showed the most promise, we decided to investigate this case further and see if making the priors for variance components less vague would improve the results.

### *3.2.5 Case IV-A Informative priors for b, but less vague for* $\sigma^2$

Using less vague priors for the variance components looks promising. The shaded area, 17 to 28 months, is symmetric and covers approximately 97% of the distribution of shelf life estimates.

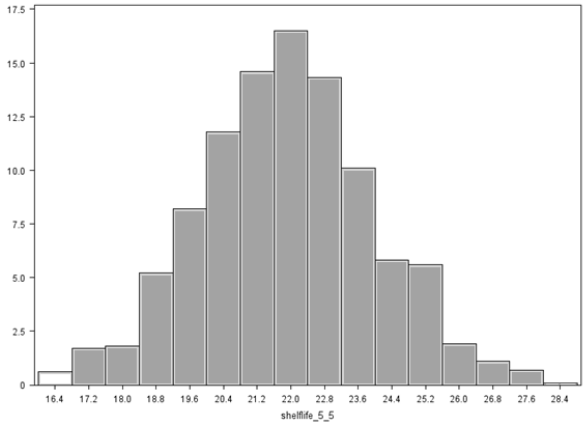



**Figure 8:** The resulting distribution of shelf life estimates using informative priors for b, but less vague priors for $\sigma^2$

The priors used for variance components that produced the distribution of shelf life given in Figure 8 above are as follows: $\sigma^2{}_0 \sim IG(0.25, 1.875), \sigma^2{}_1 \sim IG(0.25, 0.00375)$, $\sigma^2 \sim IG(0.25, 0.625)$.

To investigate what happens as we make the priors even less vague (a little more informative) we decided to look at another set of priors for the variance components.

### *3.2.6 Case IV-B Informative priors for b, but even less vague for* $\sigma^2$

Results presented in Figure 9 below were generated using the following priors for the variance components: $\sigma^2{}_0 \sim IG(0.5, 2.25)$, $\sigma^2{}_1 \sim IG(0.5, 0.00225)$, $\sigma^2 \sim IG(0.5, 0.75)$. As before, the shape and scale parameters of the inverse gamma distribution were chosen so that the mode of each distribution was equal to the parameters used to generate the "true" empirical shelf life.

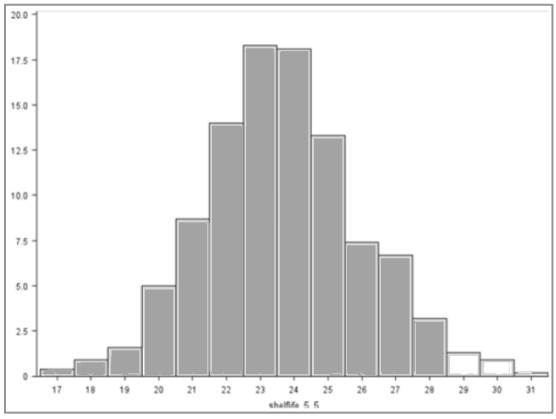



**Figure 9:** The resulting distribution of shelf life estimates using informative priors for b, but even less vague priors for $\sigma^2$

Using priors for the variance components that are even less vague (i.e. a little more informative) produces a smaller proportion of desirable estimates for shelf life compared to the previous case. Thus we will not continue our investigation in that direction and conclude that the comparison of the four different combinations of informative and vague priors showed that the highest proportion of desired estimates of shelf life is produced by a combination of informative priors on b and not too informative priors on $\sigma^2$.
To summarize a discussion in this section the combination of priors that produced the best sampling distribution of estimators of shelf life (referred to hereafter as the winning combination) is as follows: $b_0 \sim N(100, 9)$, $b_1 \sim N(-0.33, 0.05)$, ${\sigma^2}_0 \sim IG(0.25, 1.875)$, ${\sigma^2}_1 \sim IG(0.25, 0.00375)$, $\sigma^2 \sim IG(0.25, 0.625)$.

## 4. BLUP

### 4.1. Population average versus BLUP

In the previous section we investigated effect of priors on estimates of shelf life using the lower 5th confidence bound of the population average estimates for $\hat{\beta}_0$ and $\hat{\beta}_1$ produced by a random coefficient linear mixed model.
In this section we will look at the result of applying the winning combination of priors to shelf life estimates based on BLUP of the shortest-lived batch.
Figure 10 below illustrates the difference between estimates of shelf life using population average batch response over time versus shelf life estimates using BLUP of the shortest-lived batch. Dashed lines in both cases represent the lower bound of the confidence interval of the estimate used to calculate shelf life.

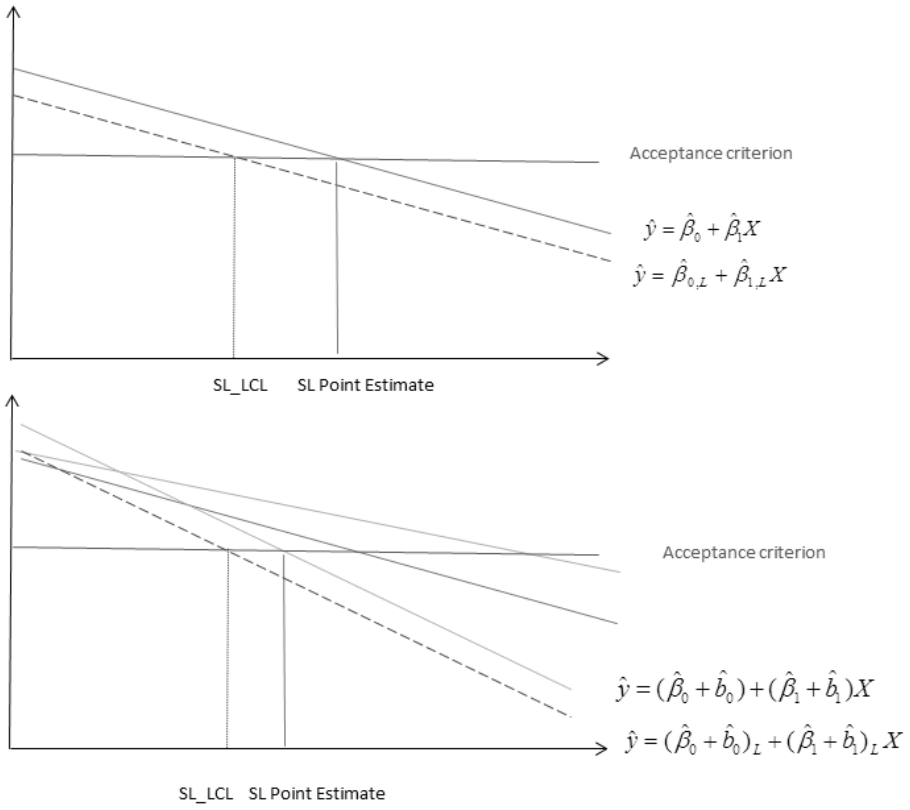



**Figure 10:** Population average versus BLUP

As mentioned, earlier previous research showed that shelf life estimates based on the lower BLUPs of the batch specific slope and intercept for the shortest-lived batch performed the best out of several other competing mixed model based approaches.
Therefore the next step of this investigation is to explore the effect of the winning combination of priors on shelf life estimates based on the BLUP of the shortest-lived batch.

Figure 11 shows the distribution of shelf lives using the winning combination of priors from section 3 based on the BLUP of the shortest-lived batch. The area shaded in grey is the desired range of shelf life estimates between 17 and 28 months. Clearly this procedure produces a large proportion of shelf life estimates that are too high. In this example the percentage of estimates that were too high is approximately 75%.

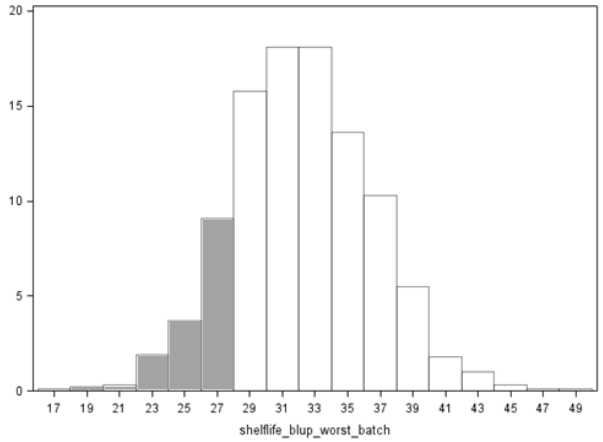



**Figure 11:** Resulting distribution of shelf life estimates using BLUP with informative priors for b, but less vague priors for $\sigma^2$

The next step of our investigation would be to use the lower confidence bound of BLUP of the shortest-lived batch.

### 4.1.2 Lower BLUP (P5)

Figure 12 shows the distribution of the estimated shelf life based on the lower $5^{th}$ percentile of the BLUP of the shortest-lived batch.
As we can see this procedure produces a large proportion of shelf life estimates that are too low. The shaded area represents the desired range of shelf life between 17 and 28 months.

From this result we can see that the lower bound that we should use for the BLUP of the shortest-lived batch is actually somewhere between the $5^{th}$ and $50^{th}$ percentiles.

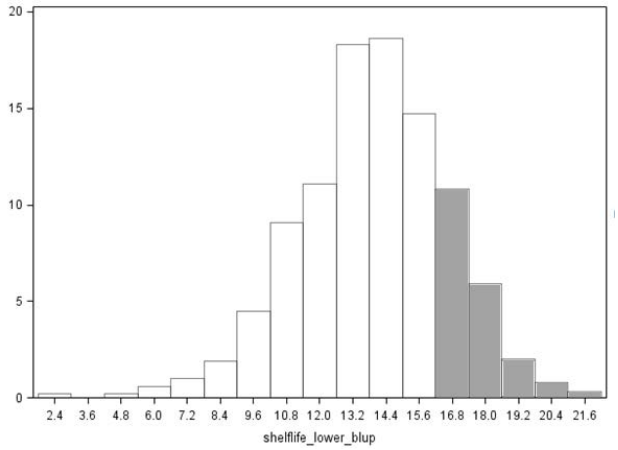



**Figure 12:** The resulting distribution of shelf life estimates using lower 5th percentile of BLUP with informative priors for b, but less vague priors for $\sigma^2$

### 4.1.2 P23

Continued trial and error investigation revealed that using the lower 23rd percentile of the BLUPs of the shortest-lived batch to estimate shelf life produced desirable results.
The shaded area on Figure 13 represents the desired range of shelf life between 17 and 28 months. We can see that shaded area covers approximately 90% of distribution.

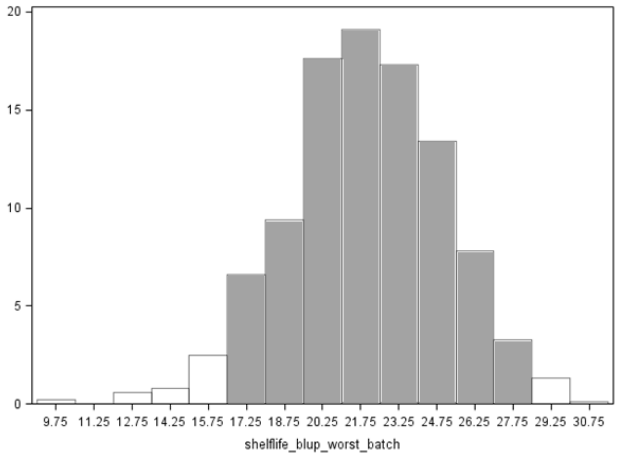



**Figure 13:** The resulting distribution of shelf life estimates using lower 23rd percentile of BLUP with informative priors for b, but less vague priors for $\sigma^2$

Further justification of usage of 23rd percentile is needed, but results show promise.

## 5. Summary and Conclusions

Using Bayesian estimated BLUPs for shelf life estimation shows promise. Informative priors that are not too informative demonstrate the best behavior. BLUPs of the shortest-lived batch to estimate shelf life based on the posterior median produce estimates that are too high, while using the lower 5th percentile of the BLUPs of the shortest-lived batch to

estimate shelf life produces estimates that are too low. Using lower 23-25$^{th}$ percentile of the BLUPs of the shortest-lived batch to estimate shelf life seems to be just right. Why it yielded the best result needs to be further investigated.

## Acknowledgements

We would like to thank Dr. Michelle Quinlan for her helpful advice, assistance and willingness to allow us to use her illustrations.

## References

J. Schwenke. "Current practices in Shelf Life Estimation," 2010, *JSM presentation.*

M. Quinlan, W. Stroup, J .Christopher & J. Schwenke. "On the distribution of batch shelf lives," 2013, *Journal of Biopharmaceutical Statistics, 23(4):897-920.*

M. Quinlan, W. Stroup, J. Schwenke, & J. Christopher. "Evaluating the performance of the ICH guidelines for Shelf life estimation," 2013, *Journal of Biopharmaceutical Statistics; 23(4):881-96.*

W. Stroup, M. Quinlan, "Statistical Consideration for Stability and the Estimating Shelf Life." Chapter 21 in *Non-Clinical Statistics for Pharmaceutical and Biotechnology Industries*, in press, to appear Fall 2015.