Supplementary material for the Biometrics paper “A Semiparametric Modeling Framework for Potential Bio-Marker Discovery and the Development of Metabonomic Profiles”

We have presented all the model comparison tables and those figures which are not covered in the main paper in the supplementary section. All model selection criteria and their statistical justification are also described in detail. We hope this additional information will prove useful and increase readability of the main paper considerably.

1. Statistical Description of Different Model Comparison Criteria

1.1 CPO Based Criterion

The CPO statistics is a very useful tool in statistical literature, first defined by Geisser (Geisser, 1980) as a diagnostic measure to detect observations discrepant from a given model. For detailed discussion on this please see (Gelfand, Dey and Chang, 1992) and (Sinha and Dey, 1997). In our situation where we have $r$ replicates from a single subject, the whole data set for a subject is denoted as $D = \bigcup_{i=1}^{r} D^i$. The CPO for the $k$-th m/z value in the $i$-th replicate is defined as,

$$CPO_k^i = \int f(t_{k_i}|D) \pi(\Theta|\bigcup_{j=1,j\neq i}^{r} D^j \cup D_{-k}^i) d\Theta,$$

where $D_{-k}^i = \{d, \{t_{nj}^{(i)}\}_{j=1,j\neq i}^{r_n}\}$ A close expression of equation (1) is not possible to obtain; however a Monte Carlo approximation given by (Gelfand, Dey and Chang, 1992) for the $CPO_k^i$ is,

$$\widehat{CPO}_k^i = \left( \frac{1}{G} \sum_{g=1}^{G} \frac{1}{f(t_{k_i}^{(g)}|\Theta^{(g)})} \right)^{-1},$$

where $\Theta^{(g)}$ denotes the parameter samples at the $g$-th, $g = 1, 2,..., G$, iteration of the Gibbs sampler. A large CPO is indicative of the better fit at that specific point. Pseudo marginal is a useful summary statistic based on the sum of the logarithm of the CPO of the individual observation is defined as

$$B^* = \sum_{i=1}^{r} \sum_{k=t_{1_i}}^{l_{n_i}} \log \left(\widehat{CPO}_k^i\right).$$

Models with greater $B^*$ values will indicate better for the specific subject profile.

1.2 Bayes Factor Based Criterion

The MCMC based approach for calculating Bayes factors was developed by (Chib, 1995), and discussed earlier by (Kass and Raftery, 1995). Given the $k$-th model, (Chib, 1995), has defined the log of Basic Marginal Likelihood Identity (BMI) as

$$\log(m_k(t)) = \log f_k(t|\Theta_k) + \log \pi(\Theta_k) - \log \hat{\pi}(\Theta_k|D),$$

and the estimated Bayes factor for two rival models indexed by $k$ and $k'$ as,
\[ B_{kk'} = \exp \{ \log(m_k(t)) - \log(m_{k'}(t)) \}. \]

We would like to take importance weighted marginal density estimation (IWMDE) based approach suggested by (Chen, 1994) to estimate \( m_k(t) \) directly. Let \( \Theta^{(g)} \) denote the parameter samples at the \( g \)-th, \( g = 1, 2, \ldots, G \), iteration of the Gibbs sampler from posterior distribution. Then IWMDE yields a consistent estimator for the \( \hat{m_k}(t) \) as:

\[
\hat{m_k}(t) = \left( \frac{1}{G} \sum_{g=1}^{G} w(\Theta^{(g)}) f(t | \Theta^{(g)}) \pi(\Theta^{(g)}) \right)^{-1},
\]

where \( w(\Theta^{(g)}) \) is a weighted density function with support \( \Omega_w \subset \Omega_{\pi(D)} \). A good \( w \) should be chosen such that it should have similar shape to the conditional marginal density, which is often unknown. However as pointed out in (Chen, 1994) the rough shape of the conditional marginal distribution can be approximately calculated through MCMC samples drawn from \( \pi(\Theta | D) \). For our purpose we have chosen a \( p \)-dimensional multivariate normal \( \mathcal{N}(\mu, \Sigma) \) as a joint importance sampling density for \( w \). We have chosen posterior mean and negative Hessian matrix to evaluate \( \mu \) and \( \Sigma \). Once the \( B_{kk'} \) has been estimated, if it turns out to be greater than one it suggests that data supports \( k \)-th model over the \( k' \)-th model and vice versa.

1.3 Bayesian Information Criterion

Bayes information criterion is a model comparison criterion based on marginal likelihoods. There are various approximation of this criterion exists in the literature. However we would like to use the approximation achieved by (Schwartz, 1978), which for a model with \( p \) parameters and \( n \) sample observation is given by

\[
BIC = -2 \log f(t | \hat{\Theta}) + p \log(n).
\]

The first term of BIC is just involving the likelihood and the second term is a penalty for model complexity. Model with lower BIC score is preferable. Once calculated we can take the difference between BIC's of the two models to get approximate Bayes factor. We would like to refer (Dudley and Haughton, 1997) for further discussion on BIC.

1.4 Mean Square Error Based Approach

Unlike survival analysis domain, in metabonomics domain, primary interest is in \( \delta_i \), which in turn depends on the intensity function \( h(t) \). This is an estimable quantity for all subjects. The whole profile analysis in the current approach works through modeling \( h(t) \). For a specific subject the mean square error (MSE) for \( m \)-th model is given by

\[
MSE(m) = \frac{1}{N} \sum_{i=1}^{r} \sum_{t=t_{k_i}}^{t_{k_{i+1}}} (d(t) - \hat{h}(t))^2 = \frac{1}{N} \sum_{i=1}^{r} \sum_{t=t_{k_i}}^{t_{k_{i+1}}} (d(t) - \hat{h}(t)) \cdot n_i \cdot n_i^2, \quad \text{where} \quad N = \sum_{i=1}^{r} \sum_{t=t_{k_i}}^{t_{k_{i+1}}} 1.
\]

A model with minimum MSE is chosen to be the best model in mean square sense.

1.5 A Cross Validation Measure
A common cross validation criterion is to leave-one-out estimator obtained by evaluating the model parameters with \( k \)-th observation removed. Let the \( k \)-th m/z value be \( t_k \) and its corresponding relative intensity \( d_k \) is being removed from the \( i \)-th replicate of a subject to obtain \( D_{i-k}^i \) as in the subsection 1.1. The available information is denoted as \( D_{(k)}^i = \bigcup_{j=1,j \neq i}^r D_j^i \cup D_{i-k}^i \). We define \( \hat{h}(t_k | D_{(k)}^i) \) as the conditional predictive intensity function (CPIF). However estimation of this quantity poses additional challenge. It is computationally infeasible even for a single subject, to estimate \( \hat{h}(t_k | D_{(k)}^i) \) for thousands of mass spectrometry data point. However through Monte Carlo sample obtained earlier, combined with numerical integration it is possible to estimate \( \hat{h}(t_k | D_{(k)}^i) \). It is easy to note that for any m/z value or \( t \) following holds

\[
\hat{h}(t \mid D_{(k)}^i) = \frac{f(t \mid D_{(k)}^i)}{1 - F(t \mid D_{(k)}^i)}.
\]

The estimation of \( f(t \mid D_{(k)}^i) \) is straightforward as discussed in equation (2). To approximately estimate \( F(t \mid D_{(k)}^i) \) we need to perform an additional numerical integration. For any \( t \in [0, \infty) \) we choose \( k \) previously available points \( t_{0} = 0, t_{1}, t_{2}, \cdots, t_{k} = t \in [0, \infty) \). For each \( t_{l} \), \((l = 1, \cdots, k)\) we calculate \( f(t_{l} \mid D_{(k)}^i) \) through equation (2). A numerical estimate of \( F(t_{l} \mid D_{(k)}^i) \) is obtained as

\[
F(t \mid D_{(k)}^i) \approx \sum_{l=1}^{k} \frac{1}{2} \left[ f(t_{l} \mid D_{(k)}^i) + f(t_{(l-1)} \mid D_{(k)}^i) \right] \times d_{l}, \text{ where } d_{l} = t_{l} - t_{(l-1)},
\]

Above works for any \( t = t_{k} \). We would like to note at \( t_{0} (= 0) \) there is no detection of ions, hence \( d_{1} = 0 \). Once we obtain \( \hat{h}(t \mid D_{(k)}^i) \) we can define a MSE based cross validation criterion

\[
CV(m) = \frac{1}{N} \sum_{i=1}^{r} \sum_{t=t_{l}}^{t_{n}} \left( d_{t} - \hat{h}(t \mid D_{(k)}^i) \cdot n_{t}^i \right)^2.
\]

Even each of the \( r_{k_i} = \left( d_{t} - \hat{h}(t \mid D_{(k)}^i) \cdot n_{t}^i \right) \) serves as a measure of model discrepancy at the \( t \)-th data point for \( i = 1, 2, \cdots, r \), and \( t = t_{1}, t_{2}, \cdots, t_{n} \). We have reported root mean square error (RMSE) value for this cross validation measure.

2. Additional SRIF and PRA plots

We provided SRIF plot for patient P-DG51 in the main paper. Figure 1-3 represents SRIF for plots of three additional patients. The main point to note that for patients SRIF curve is considerably outside the HPD credible interval obtained from the control subjects. Interestingly for patient P-DG 31, replicate 1, (Figure 2) the SRIF plot is almost within the credible interval.
However for other replicates a large deviation is observed so that we can comfortably assign patient status. This actually strengthens the necessity of replication in MS experiment.

**Figure 1:** SRIF plot for patient P-DG12. The large deviation from the 95% confidence band is an indication of irregular metabolic behavior.

**Figure 2:** SRIF plot for patient P-DG31. The large deviation from the 95% confidence band is an indication of irregular metabolic behavior.

**Figure 3:** SRIF plot for patient P-DG41. The large deviation from the 95% confidence band is an indication of irregular metabolic behavior.
Figure 4: PRA plot for patient P-DG12. Those m/z values falling outside of 95% credible interval are potential biomarkers.

Figure 5: PRA plot for patient P-DG31. Those m/z values falling outside of 95% credible interval are potential biomarkers.

Figure 6: PRA plot for patient P-DG41. Those m/z values falling outside of 95% credible interval are potential biomarkers.

As discussed in the main paper SRIF and PRA plot has two distinct purposes namely, 1> overall subject diagnostic (more global in m/z scale) 2> potential biomarker identification (more localized). Hence in SRIF a broad picture of subjects well being is captured through the plot at
much higher resolution. For PRA, we concentrate more on the finer scale to determine the list of m/z values falling outside HPD credible interval obtained from the controls. We hypothesized that for subjects with some disorder (trauma here) we will see much outlying m/z value, which is presented at PRA plot for three patients in figure 4-6. This will not be the case for a normal individual (please see figure 5 of the original paper). It is important to note that prior to SRIF and PRA based analysis, all the subject specific medications (for both control and patient) should be eliminated otherwise this may lead us to false discovery of medications as biomarker.

3. Additional Model Comparison Results

Tables 1-3 show different model parameter estimates, standard deviations and their HPD credible intervals. Notably we are not plugging the posterior estimates of different parameters directly during estimation stage. Details of pooling control subjects are being explained latter on. Tables 4-8 represents model comparison results for three models under different model selection criteria. Within the domain of these model choices we exhaustively searched for the best model. An additional flexibility analysis is also reported latter to show that by increasing the number of mixture parameters, we will not gain much in terms capturing local features.

*Estimates of the parameters for 12 subjects for Model 1*

| Subject ID | Model 1: Extreme Value | Parameter Estimate | Posterior mean, S.D. and 95% H.P.D. Credible Interval |
|------------|------------------------|--------------------|------------------------------------------------------|
| P-DC1      |                        | 0.243              | 0.155 0.179 0.154 0.205 3.005 0.152 0.059 0.156 |
|            |                        | (0.127, 0.354)     | (0.077, 0.301) (0.052, 0.304) (0.07, 0.312) (0.04, 0.374) (2.135, 4.144) (0.11, 0.62) |
| P-DG3      |                        | 0.239              | 0.179 0.152 0.155 0.213 3.005 0.154 0.075 0.2 |
|            |                        | (0.115, 0.354)     | (0.063, 0.295) (0.074, 0.302) (0.07, 0.293) (0.06, 0.354) (2.2, 4.16) (0.11, 0.59) |
| P-DG12     |                        | 0.24               | 0.152 0.118 0.159 0.211 3.027 0.176 0.079 0.215 |
|            |                        | (0.112, 0.346)     | (0.077, 0.303) (0.079, 0.304) (0.065, 0.295) (0.051, 0.373) (2.19, 4.11) (0.12, 0.62) |
| P-DG31     |                        | 0.23               | 0.151 0.179 0.156 0.212 3.012 0.176 0.075 0.204 |
|            |                        | (0.119, 0.36)      | (0.073, 0.305) (0.059, 0.254) (0.052, 0.305) (0.067, 0.376) (2.497, 4.09) (0.13, 0.53) |
| P-DG41     |                        | 0.242              | 0.153 0.179 0.156 0.209 3.009 0.191 0.09 0.265 |
|            |                        | (0.129, 0.36)      | (0.05, 0.302) (0.075, 0.293) (0.065, 0.309) (0.069, 0.356) (2.02, 4.36) (0.10, 0.71) |
| P-DG51     |                        | 0.197              | 0.186 0.186 0.196 0.21 2.354 0.276 0.101 0.21 |
|            |                        | (0.091, 0.302)     | (0.093, 0.307) (0.101, 0.301) (0.09, 0.295) (0.024, 0.402) (0.025, 0.56) (0.05, 0.65) |
| C-DG32     |                        | 0.244              | 0.154 0.153 0.154 0.204 2.9 0.18 0.073 0.209 |
|            |                        | (0.124, 0.367)     | (0.07, 0.295) (0.077, 0.305) (0.073, 0.212) (0.062, 0.304) (1.73, 3.55) (0.09, 0.65) |
| C-DG63     |                        | 0.244              | 0.154 0.153 0.154 0.204 2.9 0.18 0.073 0.209 |
|            |                        | (0.124, 0.367)     | (0.07, 0.295) (0.077, 0.305) (0.073, 0.212) (0.062, 0.304) (1.73, 3.55) (0.09, 0.65) |
| C-DG84     |                        | 0.244              | 0.154 0.153 0.154 0.204 2.9 0.18 0.073 0.209 |
|            |                        | (0.124, 0.367)     | (0.07, 0.295) (0.077, 0.305) (0.073, 0.212) (0.062, 0.304) (1.73, 3.55) (0.09, 0.65) |
| C-DG96     |                        | 0.244              | 0.154 0.153 0.154 0.204 2.9 0.18 0.073 0.209 |
|            |                        | (0.124, 0.367)     | (0.07, 0.295) (0.077, 0.305) (0.073, 0.212) (0.062, 0.304) (1.73, 3.55) (0.09, 0.65) |
| C-DG70     |                        | 0.244              | 0.154 0.153 0.154 0.204 2.9 0.18 0.073 0.209 |
|            |                        | (0.124, 0.367)     | (0.07, 0.295) (0.077, 0.305) (0.073, 0.212) (0.062, 0.304) (1.73, 3.55) (0.09, 0.65) |
| C-DG72     |                        | 0.244              | 0.154 0.153 0.154 0.204 2.9 0.18 0.073 0.209 |
|            |                        | (0.124, 0.367)     | (0.07, 0.295) (0.077, 0.305) (0.073, 0.212) (0.062, 0.304) (1.73, 3.55) (0.09, 0.65) |

---

Table 1
### Table 2

**Estimates of the parameters for 12 subjects for Model 2**

| Subject ID | Parameter Estimate | Posterior mean, S.D. and 95% H.P.D. Credible Interval |
|------------|--------------------|--------------------------------------------------------|
|  | $\eta_1$ | $\eta_2$ | $\eta_3$ | $\eta_4$ | $\eta_5$ | $\alpha$ | $\beta$ |
| **P-DG1** | 0.139 | 0.07 | 0.036 | 0.019 | 0.73 | 2.51 | 0.127 |
| & (0.043, 0.23, 0.021, 0.198, 0.0802) | & (1.984, 2.94, 0.068, 0.403) |
| **P-DG3** | 0.199 | 0.103 | 0.059 | 0.044 | 0.594 | 2.65 | 0.105 |
| & (0.023, 0.23, 0.021, 0.198, 0.0802) | & (2.183, 2.97, 0.059, 0.232) |
| **P-DG12** | 0.135 | 0.073 | 0.052 | 0.019 | 0.72 | 2.51 | 0.109 |
| & (0.021, 0.23, 0.026, 0.198, 0.0802) | & (2.103, 2.79, 0.062, 0.327) |
| **P-DG31** | 0.153 | 0.127 | 0.119 | 0.122 | 0.165 | 2.40 | 0.129 |
| & (0.026, 0.23, 0.021, 0.198, 0.0802) | & (1.93, 2.921, 0.074, 0.39) |
| **P-DG41** | 0.179 | 0.093 | 0.052 | 0.034 | 0.641 | 2.56 | 0.082 |
| & (0.023, 0.23, 0.021, 0.198, 0.0802) | & (2.16, 2.78, 0.04, 0.216) |
| **P-DG51** | 0.154 | 0.065 | 0.041 | 0.027 | 0.71 | 2.43 | 0.075 |
| & (0.023, 0.23, 0.021, 0.198, 0.0802) | & (-3.38, 2.19, 0.01, 0.251) |
| **C-DG62** | 0.134 | 0.071 | 0.03 | 0.021 | 0.741 | 2.43 | 0.07 |
| & (0.023, 0.23, 0.021, 0.198, 0.0802) | & (2.27, 0.06, 0.33) |
| **C-DG63** | 0.139 | 0.067 | 0.042 | 0.018 | 0.73 | 2.46 | 1.01 |
| & (0.023, 0.23, 0.021, 0.198, 0.0802) | & (1.93, 2.85, 0.071, 1.41) |
| **C-DG64** | 0.135 | 0.074 | 0.037 | 0.019 | 0.73 | 2.38 | 0.142 |
| & (0.023, 0.23, 0.021, 0.198, 0.0802) | & (1.93, 2.93, 0.081, 0.474) |
| **C-DG66** | 0.143 | 0.067 | 0.036 | 0.018 | 0.74 | 2.37 | 0.115 |
| & (0.023, 0.23, 0.021, 0.198, 0.0802) | & (1.75, 2.79, 0.081, 0.467) |
| **C-DG70** | 0.143 | 0.075 | 0.039 | 0.019 | 0.72 | 2.40 | 0.03 |
| & (0.023, 0.23, 0.021, 0.198, 0.0802) | & (2.04, 2.67, 0.054, 0.243) |
| **C-DG72** | 0.142 | 0.07 | 0.051 | 0.024 | 0.72 | 2.48 | 0.128 |
| & (0.023, 0.23, 0.021, 0.198, 0.0802) | & (2.02, 2.97, 0.072, 0.491) |
## Estimates of the parameters for 12 subjects for Model 3

| Subject ID | \( \eta_1 \) | \( \eta_2 \) | \( \eta_3 \) | \( \eta_4 \) | \( \eta_5 \) | \( \alpha \) | \( \beta \) |
|------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| P-DG1      | 0.300         | 0.184         | 0.177         | 0.179         | 0.149         | 3.356         | 34.8          |
|            | (0.124,0.417) | (0.697,0.292) | (0.086,0.283) | (0.086,0.282) | (0.044)      | (1.787)       | (74.48)       |
| P-DG3      | 0.338         | 0.170         | 0.131         | 0.131         | 0.140         | 3.25          | 7.236         |
|            | (0.680,0.427) | (0.681,0.273) | (0.093,0.283) | (0.094,0.283) | (0.041)      | (1.870)       | (40.955)      |
| P-DG12     | 0.300         | 0.170         | 0.18          | 0.177         | 0.152         | 3.31          | 82.36         |
|            | (0.121,0.417) | (0.686,0.274) | (0.092,0.276) | (0.092,0.276) | (0.042)      | (1.957)       | (94.327)      |
| P-DG31     | 0.315         | 0.184         | 0.18          | 0.18          | 0.159         | 3.38          | 54.873        |
|            | (0.133,0.417) | (0.660,0.283) | (0.083,0.284) | (0.086,0.283) | (0.040)      | (1.793)       | (57.28)       |
| P-DG41     | 0.317         | 0.134         | 0.175         | 0.178         | 0.145         | 3.349         | 34.513        |
|            | (0.137,0.43)  | (0.095,0.282) | (0.082,0.267) | (0.081,0.267) | (0.046)      | (1.584)       | (68.454)      |
| P-DG51     | 0.197         | 0.195         | 0.196         | 0.2           | 0.211         | 2.41          | 793.142       |
|            | (0.095,0.307) | (0.093,0.304) | (0.095,0.31)  | (0.095,0.31)  | (0.0105)     | (2.005)       | (752)         |

Table 3
Log of Pseudo Marginal Likelihood: Model 2 is significantly better than the other two models. However for subject P-DG1 and C-DG72 Model 1 is performing better.

| Subject ID | Log of Pseudo Marginal Likelihood |
|------------|-----------------------------------|
|            | Model 1  | Model 2  | Model 3  | Best Model |
| P-DG1      | -11411   | -44215   | -2.24E+12| Model 1     |
| P-DG3      | -91093   | -75545   | -2.00E+12| Model 2     |
| P-DG12     | -17905   | -15302   | -4.11E+12| Model 2     |
| P-DG31     | -9427    | -7763    | -1.57E+12| Model 2     |
| P-DG41     | -27345   | -22505   | -4.00E+14| Model 2     |
| P-DG51     | -63000   | -62000   | -4.07E+12| Model 2     |
| C-DG62     | -11303   | -9209    | -2.02E+12| Model 2     |
| C-DG63     | -13440   | -11593   | -4.01E+12| Model 2     |
| C-DG64     | -11047   | -9566    | -4.02E+12| Model 2     |
| C-DG66     | -7204    | -6497    | -1.97E+11| Model 2     |
| C-DG70     | -15505   | -14735   | -3.31E+12| Model 2     |
| C-DG72     | -20253   | -35407   | -4.97E+12| Model 1     |

Table 4

Bayes’ Factor: Model 2 is significantly better than the other two models.

| Subject ID | Logarithm of Bayes’ Factor |
|------------|----------------------------|
|            | $B_{12}$  | $B_{13}$  | $B_{23}$  | IWMDE      | Best Model |
| P-DG1      | -2883.224| 319034.218| 319034.218| Model 2     |
| P-DG3      | -6845.901| 377396.254| 384242.155| Model 2     |
| P-DG12     | -3662.78 | 12385.733 | 16048.513 | Model 2     |
| P-DG31     | -2099.623| 238165.032| 240264.655| Model 2     |
| P-DG41     | -6937.871| 230929.214| 237867.085| Model 2     |
| P-DG51     | 1767.618 | 1052701.959| 1050934.341| Model 2     |
| C-DG62     | -2804.496| 302312.22  | 305116.716| Model 2     |
| C-DG63     | -2094.925| 228732.135| 230827.06 | Model 2     |
| C-DG64     | -1696.396| 238717.695| 240414.091| Model 2     |
| C-DG66     | -1263.226| 277006.82  | 278270.046| Model 2     |
| C-DG70     | -4477.42 | 459893.596 | 464371.016| Model 2     |
| C-DG72     | -2929.461| 207293.392 | 210222.853| Model 2     |

Table 5
Bayesian Information Criterion: Model 2 is significantly better than the other two models.

Model 1: Extreme Value, Model 2: Double Exponential, Model 3: Normal

| Subject ID | Bayesian Information Criterion (BIC) | Best Model |
|------------|-------------------------------------|------------|
| P-DG1      | Model 1: 51015.14, Model 2: 29571.17, Model 3: 126501.2 | Model 2     |
| P-DG3      | Model 1: 225992.64, Model 2: 191259.5, Model 3: 94067.46 | Model 3     |
| P-DG12     | Model 1: 45044.66, Model 2: 35004.11, Model 3: 51206.8 | Model 2     |
| P-DG31     | Model 1: 22695.55, Model 2: 15792, Model 3: 30673.2 | Model 2     |
| P-DG41     | Model 1: 66365.01, Model 2: 52357.9, Model 3: 65950.52 | Model 2     |
| P-DG51     | Model 1: 101219.4, Model 2: 97565.3, Model 3: 94144.42 | Model 3     |

C-DG62  | Model 1: 25626.55, Model 2: 20119.36, Model 3: 25704 | Model 2     |
| C-DG63  | Model 1: 30369.75, Model 2: 26365.52, Model 3: 34959.76 | Model 2     |
| C-DG64  | Model 1: 25551.26, Model 2: 22325.29, Model 3: 31425.52 | Model 2     |
| C-DG66  | Model 1: 15095.42, Model 2: 15749.02, Model 3: 29624.6 | Model 2     |
| C-DG70  | Model 1: 46622.19, Model 2: 36429.49, Model 3: 50577.29 | Model 2     |
| C-DG72  | Model 1: 43376.7, Model 2: 36550.5, Model 3: 47263.23 | Model 2     |

Table 6

Mean Square Error: Model 1 is better for some of the subjects while Model 2 is performing better for some other subjects.

Model 1: Extreme Value, Model 2: Double Exponential, Model 3: Normal

| Subject ID | Mean Square Error (MSE) | Best Model |
|------------|-------------------------|------------|
| P-DG1      | Model 1: 2752208.9, Model 2: 66.61, Model 3: almost $\infty$ | Model 2     |
| P-DG3      | Model 1: 179.59, Model 2: 306.465, Model 3: 506.76 | Model 1     |
| P-DG12     | Model 1: 319.84, Model 2: 34.53, Model 3: 1292.7 | Model 2     |
| P-DG31     | Model 1: 17.7, Model 2: 25.03, Model 3: 373.6 | Model 1     |
| P-DG41     | Model 1: 58.14, Model 2: 131.5, Model 3: 5980.3 | Model 1     |
| P-DG51     | Model 1: 9.924E+14, Model 2: 7.084E+11, Model 3: almost $\infty$ | Model 2     |

C-DG62  | Model 1: 18.23, Model 2: 33.97, Model 3: 7024.6 | Model 1     |
| C-DG63  | Model 1: 18.84, Model 2: 37.7, Model 3: 3269.6 | Model 1     |
| C-DG64  | Model 1: 20.68, Model 2: 28.02, Model 3: 1435.5 | Model 1     |
| C-DG66  | Model 1: 14.6, Model 2: 20.15, Model 3: 20024.6 | Model 1     |
| C-DG70  | Model 1: 36.32, Model 2: 25.93, Model 3: 574.2 | Model 2     |
| C-DG72  | Model 1: 51.09, Model 2: 65.84, Model 3: 2.082E+10 | Model 1     |

Table 7
Cross Validation based RMSE: Model 1 is better for some of the subjects while Model 2 is performing better for some other subjects.

Model 1: Extreme Value, Model 2: Double Exponential, Model 3: Normal

| Subject ID   | Root Mean Square Error (RMSE) |
|--------------|-------------------------------|
|              | Model 1 | Model 2 | Model 3 | Best Model |
| P-DG1        | 60882.02 | 180.329 | almost ∞ | Model 2    |
| P-DG3        | 20.85   | 20.84   | 631.88  | Model 2    |
| P-DG12       | 22.506  | 42.55   | 8002.52 | Model 1    |
| P-DG31       | 71.64   | 137.95  | 713.16  | Model 1    |
| P-DG41       | 35.74   | 87.26   | 9980.74 | Model 1    |
| P-DG51       | 57.15   | 57.17   | almost ∞| Model 1    |
| C-DG62       | 7.16    | 11.03   | 8025.06 | Model 1    |
| C-DG63       | 38.93   | 58.07   | 2286.13 | Model 1    |
| C-DG64       | 33.27   | 47.62   | 1435.5  | Model 1    |
| C-DG66       | 45.68   | 61.07   | 16624.52| Model 1    |
| C-DG70       | 16.69   | 16.67   | 351.98  | Model 2    |
| C-DG72       | 32.92   | 10.05   | almost ∞| Model 2    |

Table 8

4. Pooling of Control Subjects to Construct HPD

We are estimating seven parameters namely five mixture parameters ($\eta$), location ($\alpha$) and scale ($\beta$) parameter in $\tilde{H}(t)$ for each subject. These help us estimating relative intensity (hazard) function given by

$$h(t) = \frac{\frac{d}{dt} \tilde{H}(t) \sum_{l=1}^{m} \eta_l Bc(\tilde{H}(t); r_l, s_l)}{\{1 - \sum_{l=1}^{m} \eta_l IB(\tilde{H}(t); r_l, s_l)\}^2}.$$  

Likelihood for a subject is given in equation (3) of the main paper. Note that for each of the control subjects we are obtaining samples from the posterior distributions of each parameter. Under the hypothesis that control subjects are homogeneous, the posterior parameter distribution should be similar. To illustrate more consider location parameter $\alpha$ only. We obtain posterior samples from all control subjects separately and then mix them together under the hypothesis that posterior distribution of $\alpha$ is same for all controls. Notably this will not be true if underlying hypothesis of homogeneity among controls is violated. Since different trauma types are highly heterogeneous, it is a prime reason that we cannot combine patient samples. Then we repeat the same procedure for other parameters too.

Figure 7 represents SRIF plots for same subject (patient) for two different replicates. Evidently credible intervals are different and apparently may not be clear to the reader why it should be so even for the same patient across replicates. Notably credible intervals are constructed exclusively by combining controls, and of course once posterior samples are obtained they never change. However note that $h(t)$ is also a function of $t$ (proxy of the m/z value) with seven other
parameters and the location of the observed intensities ($t$) are different even between replicate to replicate. This is depicted in figure 8.

![figure 7](image1.png)

**Figure 7**

![figure 8](image2.png)

**Figure 8**

We calculate credible interval for intensity which corresponds to a patient based on a replicate at a position where we observe the intensity value. Since $t_1$ for replicate 1 is different than $t_1$ in replicate 2 this implies associated $h(t)$ value and its credible interval will be also different. In fact $t_1$ denotes the first m/z value (location) of a replicate and $t_1$ across different replicates has no straight forward relationship except for a place holder. Notably we have not used any alignment algorithm in the current study. However they ($t$ values) will be exactly same if the replicates are in perfect alignment. The PRA estimate is given by

$$\hat{\delta}_i^3 = \hat{h}(t) \times n_i^t \text{ for } t = t_1, t_2, \ldots, t_n.$$  

As $\delta_i$ is very much depends upon $h(t)$, which in turns depends on observed $t$ of a specific replicate (say $i$ -th), it is necessary to introduce a superscript-$i$ in $\delta_i$. Since we have many
posterior samples for each parameter from controls, we can obtain several estimate of $\delta_i$ for fixed $t$ and $i$-th replicate. Hence we introduced a Monte Carlo (MC) estimate of $\delta_i^t$ as

$$\hat{\delta}_i^t = n_i^t \times \frac{1}{G} \sum_{g=1}^{G} \hat{h}_g(t).$$

It is easy to construct HPD credible interval from G-many MC estimates of $\delta_i^t$ following Chen-Shao’s (JCGS, 1998) algorithm. Notably if there is no variation in the x-axis (observed m/z value) we do not need i-th subscript in $\delta_i$ consequently credible intervals will not vary across replicates.

Similar conclusion can be drawn from SRIF which is expressed as

$$S_{MC}^t = \frac{1}{G} \sum_{g=1}^{G} S_{g}^t = \frac{1}{G} \sum_{g=1}^{G} \prod_{j:t_j \leq t} (1 - \hat{h}_g(t_j)).$$

Again SRIF depends upon estimation of $h(t)$ and all the above reasoning will also be valid for difference in credible interval across replicates.

5. Some Comments About Flexibility with Five beta Mixtures

A pertinent question is whether the considered model can capture complexity in the data with only five beta mixtures. We performed several flexibility analyses by varying number of mixands at different levels. In our experiment we have tried increasing the number of components from five to ten. While more mixture components will somewhat improve the result however it comes at the expense of increasing the computational cost. We would like to report that for the subject specific global profile estimation (SRIF) five mixands is as good as ten for all three models considered here. However this is highly study specific and we recommended trying several of them in consultation with domain experts. In principle parsimonious model selection should be encouraged. Another way out is to treat the number of mixing components as an unknown quantity which needs to be estimated from the data. However this will come at the increasing cost of computational demand.

Additional References

Geisser, S. (1980). Discussion on sampling and bayes' inference in scientific modeling and robustness (by g.e.p. box). J. R. Statist. Soc. A 143, 416-417.

Gelfand, A. E., Dey, D. K. and Chang, H. (1992). Model determination using predictive distribution with implementation via sampling based methods. In Bernardo, J. M., Berger, J. O., Dawid, A. P. and Smith, A. F. M., editors, Bayesian Statistics 4, pages 147-167. Oxford University Press.

Sinha, D. and Dey, D. K. (1997). Semiparametric bayesian analysis of survival data. Journal of the American Statistical Association 92, 1195-1212.

Chib, S. (1995). Marginal likelihood from the gibbs output. Journal of the American Statistical Association 90, 1313-1321.
Kass, R. and Raftery, A. E. (1995). Bayes factor and model uncertainty. Journal of the American Statistical Association 90, 773-795.

Chen, M. H. (1994). Importance-weighted marginal bayesian posterior density estimation. Journal of the American Statistical Association 89, 818-824.

Schwartz, G. (1978). Estimating the dimension of a model. Ann. Statist. 6, 461-466.