Supplement: Bayesian Nonparametric Estimation for Dynamic Treatment Regimes with Sequential Transition Times

A: Details of MCMC for Fitting DDP-GP Model

Summary of Model

$$p(y_i^k \mid \boldsymbol{x}_i^k, F^k) = F^k(y_i^k \mid \boldsymbol{x}_i^k)$$
$$F^k \sim \text{DDP-GP}\left\{\{\mu_h^k\}, C^k; \alpha^k, \{\boldsymbol{\beta}_h^k\}, \sigma^k\right\} \tag{1}$$

$$F^k(y \mid \boldsymbol{x}^k) = \sum_{h=0}^{\infty} w_h^k \, N(y;\ \theta_h^k(\boldsymbol{x}^k),\ \sigma^k). \tag{2}$$

$$\{\theta_h^k(\boldsymbol{x}^k)\} \sim\ GP(\mu_h^k(\boldsymbol{x}^k), C^k(\boldsymbol{x}^k)). \qquad h = 1, 2, ...$$

$$\mu_h^k(\boldsymbol{x}_i^k) = \boldsymbol{x}_i^k \boldsymbol{\beta}_h^k.$$

$k = 1, \ldots, n_{\text{trans}}$. We complete the model construction by assuming $\boldsymbol{\beta}_h^k \sim N(\boldsymbol{\beta}_0^k, \Sigma_0^k)$, $(\sigma^k)^{-2} \overset{\text{i.i.d.}}{\sim} \text{Ga}(\lambda_1, \lambda_2)$ and $\alpha^k \overset{\text{i.i.d.}}{\sim} \text{Ga}(\lambda_3, \lambda_4)$.

Posterior Computation

To evaluate the posterior in a DDP-GP model, we first marginalize (1) analytically with respect to the random probability measures $F^k(\cdot|\boldsymbol{x}^k)$. The form is not obvious from the earlier definition. Temporarily suppress the superscripted transition index k. Consider generating a sample $(Y_1, \boldsymbol{x}_1)$, $\cdots$, $(Y_n, \boldsymbol{x}_n)$ by first sampling from a covariate distribution, $p(\boldsymbol{x})$, and then from a conditional transition time distribution, $F(\cdot|\boldsymbol{x})$. We rewrite (2) as a hierarchical model with a new latent indicator variable γ_i for the normal mixture summand index h,

$$(Y_i \mid \gamma_i = h,\ \ \boldsymbol{x}_i) \sim N(\theta_h(\boldsymbol{x}_i), \sigma^2) \ \text{ and } \ p(\gamma_i = h) = w_h, \tag{3}$$

for $i = 1, \cdots, n$. Let $\widetilde{\theta}_i(\cdot) = \theta_{\gamma_i}(\cdot)$ denote a realization of the stochastic process selected by γ_i. Next, we re-index the $\theta_h(\cdot)$ such that $\sum_{i=1}^n I(\gamma_i = h) \geq 1$ for $h = 1, \ldots, H$. That is, we let $h = 1, \ldots, H$ index the realizations $\theta_h(\cdot)$ that are selected by some of the γ_i's, so

that $\{\theta_1, \cdots, \theta_H\}$ are the unique values of the n realizations $\{\widetilde{\theta}_i,\ i = 1, \ldots, n\}$. If clusters of patients are defined as $S_h = \{i : \ \widetilde{\theta}_i = \theta_h\}$, then the γ_i's are interpreted as cluster membership indicators. Posterior simulation makes use of these indicators and the vectors $\boldsymbol{\theta}_h = (\theta_h(\boldsymbol{x}_1), \ldots, \theta_h(\boldsymbol{x}_n))$. After marginalization with respect to F_x, we are left with the marginal model for $\{\gamma_i,\ \boldsymbol{\theta}_h(\boldsymbol{x}_i);\ i = 1, \ldots, n,\ h = 1, \ldots, H\}$.

For each transition k, we update parameters using finite DP algorithm as follows. Denote $\#\{i : \gamma_i^k = h\} = n_h^k$.

- Update σ^k

$$(\sigma^k)^2 \mid \cdot \sim \text{Inverse Gamma}(\lambda_1 + \frac{n^k}{2}, \lambda_2 + \frac{\sum_{h=1}^{H}\sum_{\gamma_i^k=h}(y_i^k - \theta_h^k(\boldsymbol{x}_i))^2}{2}) \tag{4}$$

- Update $\boldsymbol{\theta}_h^k$

$$\begin{aligned} p(\boldsymbol{\theta}_h^k \mid \cdot) &\propto p(\boldsymbol{\theta}_h^k) \prod_{i:\gamma_i^k=h} p(y_i^k \mid \theta_h^k(\boldsymbol{x}_i^k)) \\ &\propto \exp\{-\frac{1}{2}(\boldsymbol{\theta}_h^k - \boldsymbol{X}^k\boldsymbol{\beta}_h^k)'(C^k)^{-1}(\boldsymbol{\theta}_h - \boldsymbol{X}^k\boldsymbol{\beta}_h^k)\} \times \exp\{-\frac{\sum_{i:\gamma_i^k=h}(y_i^k - \theta_h^k(\boldsymbol{x}_i))^2}{2(\sigma^k)^2}\} \\ &\sim \text{N}(((C^k)^{-1} + \frac{U'U}{(\sigma^k)^2}I)^{-1}(U'\frac{\boldsymbol{y}_h^k}{(\sigma^k)^2} + (C^k)^{-1}\boldsymbol{X}^k\boldsymbol{\beta}_h^k), ((C^k)^{-1} + \frac{U'U}{(\sigma^k)^2}I_{n^k\times n^k})^{-1}), \end{aligned}$$

where $\boldsymbol{y}_h^k = \{y_i^k, \gamma_i^k = h\}$, $I_{n^k\times n^k}$ is an $n^k \times n^k$ identity matrix, $\boldsymbol{X}^k$ is an $n^k \times M^k$ matrix with the covariates $\boldsymbol{x}_i^k$ of the i-th patient in row i. U is a $n_h^k \times n^k$ matrix: if patient i is the j-th element of $\gamma_i^k = r$, then $U_{ji} = 1$. All other elements are 0.

- Update $\boldsymbol{\beta}_h^k$

$$\begin{aligned} p(\boldsymbol{\beta}_h^k \mid \cdot) &\propto p(\boldsymbol{\beta}_h^k)\exp\{-\frac{1}{2}(\boldsymbol{\theta}_h^k - \boldsymbol{X}^k\boldsymbol{\beta}_h^k)'(C^k)^{-1}(\boldsymbol{\theta}_h^k - \boldsymbol{X}^k\boldsymbol{\beta}_h^k)\} \\ &\sim \text{N}(\Sigma_h^k[((\boldsymbol{X}^k)'(C^k)^{-1}\boldsymbol{\theta}_h^k + \Sigma_0^k\boldsymbol{\beta}_0^k], \Sigma_h^k), \end{aligned}$$

where $\Sigma_h^k = ((\boldsymbol{X}^k)'(C^k)^{-1}\boldsymbol{X}^k + (\Sigma_0^k)^{-1})^{-1}$.

- Update w_h^k

$$v_h^k \sim Beta(1 + n_h^k, \alpha^k + \sum_{j>h} n_j^k),$$

where $n_h^k = \sum_{i=1}^{n_k} I(\gamma_i^k = h)$ is the number of observations such that $\gamma_i^k = h$. Then $w_h^k = v_h^k \prod_{j>h}(1 - v_j^k)$.

- Update γ_i^k

  - If y_i^k is not censored,

$$Pr(\gamma_i^k = h \mid \cdot) \quad \propto \quad w_h^k \int p(y_i^k \mid \theta_h^k(\boldsymbol{x}_i)) p(\theta_h^k(\boldsymbol{x}_i^k) \mid \theta_h^k(\boldsymbol{x}_{-i}^k)) d(\theta_h^k(\boldsymbol{x}_i^k)),$$

    where $\boldsymbol{x}_{-i}^k = \{\boldsymbol{x}_j^k : \gamma_j^k = h, j \neq i\}$.

  - If y_i^k is censored, Let

$$p_h^k(t) = \int p(y_i^k \mid \theta_h^k(\boldsymbol{x}_i)) p(\theta_h^k(\boldsymbol{x}_i^k) \mid \theta_h^k(\boldsymbol{x}_{-i}^k)) d(\theta_h^k(\boldsymbol{x}_i^k)).$$

    Then

$$Pr(\gamma_i^k = h \mid \cdot) \quad \propto \quad \int_{V_i^k}^{\infty} w_h^k p_h^k(t) dt,$$

    where V_i^k is the observed time for patient i in transition k.

- Update α^k

  Using data augmentation, we first sample an m from beta distribution beta($\alpha^k + 1, n^k$). Then we sample the new α^k value from

$$\alpha^k \sim \pi \mathrm{Ga}(\lambda_3 + H, \lambda_4 - \log(m)) + (1 - \pi)\mathrm{Ga}(\lambda_3 + H - 1, \lambda_4 - \log(m)),$$

  where $\frac{\pi}{1-\pi} = \frac{\lambda_3 + H - 1}{n^k(\lambda_4 - \log(m))}$.

B: Determining Prior Hyperparameters

As priors for $\boldsymbol{\beta}_h^k$ in (1) we assume $\boldsymbol{\beta}_h^k \sim N(\boldsymbol{\beta}_0^k, \Sigma_0^k)$ for each transition k, $h = 1, 2, \ldots$. For σ^k we assume $(\sigma^k)^{-2} \overset{\text{i.i.d.}}{\sim} \text{Ga}(\lambda_1, \lambda_2)$. And, finally, $\alpha^k \overset{\text{i.i.d.}}{\sim} \text{Ga}(\lambda_3, \lambda_4)$.

To apply the DDP-GP model, one must first determine numerical values for the fixed hyperparameters $\{\boldsymbol{\beta}_0^k, \Sigma_0^k, k = 1, 2, ...\}$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. This is a critical step. These numerical hyperparameter values must facilitate posterior computation, and they should not introduce inappropriate information into the prior that would invalidate posterior inferences. With this in mind, the hyperparameters $(\boldsymbol{\beta}_0^k, \Sigma_0^k)$ for the k^{th} transition time covariate effect distribution may be obtained via empirical Bayes by doing a preliminary fits of a lognormal distribution $Y^k = \log(T^k) \sim N(\boldsymbol{x}^k \boldsymbol{\beta}_0^k, \sigma_0^k)$ for each transition k. Similarly, we assume a diagonal matrix for Σ_0^k with the diagonal values also obtained from the preliminary fit of the lognormal distribution. Once an empirical estimate of σ^k is obtained, one can tune (λ_1, λ_2) so that the prior mean of σ^k matches the empirical estimate and the variance equals 1 or a suitably large value to ensure a vague prior. Finally, information about α^k typically is not available in practice. We use $\lambda_3 = \lambda_4 = 1$.

This approach works in practice because the parameter $\boldsymbol{\beta}_0^k$ specifies the prior mean for the mean function of the GP prior, which in turn formalizes the regression of T^k on the covariates $\boldsymbol{x}^k$, including treatment selection. The imputed treatment effects hinge on the predictive distribution under that regression. Excessive prior shrinkage could smooth away the treatment effect that is the main focus. The use of an empirical Bayes type prior in the present setting is similar to empirical Bayes priors in hierarchical models. This type of empirical Bayes approach for hyperparameter selection is commonly used when a full prior elicitation is either not possible or is impractical. Inference is not sensitive to values of the hyperparameters $\boldsymbol{\lambda}$ that determine the priors of σ^k and α^k for two reasons. First, the standard deviation σ^k is the scale of the kernel that is used to smooth the discrete random probability measure generated by the DDP prior. It is important for reporting a smooth fit, that is for display, but it is not critical for the imputed fits in our regression setting. Assuming some regularity of the posterior mean function, smoothing adds only minor corrections. Second, the total mass parameter α^k determines the number of unique clusters formed in the underlying Polya urn. However, because most clusters are small changing the prior of α^k does not significantly change the posterior predictive values that

are the basis for the proposed inference.

The conjugacy of the implied multivariate normal on $\{\theta_h^k(\boldsymbol{x}_i^k), i = 0, \ldots, n\}$ and the normal kernel in (2) greatly simplify computations, since any Markov chain Monte Carlo (MCMC) scheme for DP mixture models can be used. MacEachern and Müller (1998) and Neal (2000) described specific algorithms to implement posterior MCMC simulation in DPM models. Ishwaran and James (2001) developed alternative computational algorithms based on finite DPs, which truncated (2) after a finite number of terms.

C: Survival Time Regression Simulation

This simulation was designed to study the DDP-GP regression model by comparing inference for a survival function with the simulation truth. In this study, we did not evaluate a regime effect, but rather focused on inference for the survival curve.

For each subject, we generated T = survival time, the covariates x_1 = tumor size (0=small, 1=large) and x_2 = body weight, and x_3 = a biomarker (0=absent, 1=present). We assumed that small and large tumor sizes each had probability .50. Body weights were computed by sampling from a uniform distribution, Unif$(80, 150)$, with the covariate x_2 defined by shifting and scaling to obtain mean 0 and variance 1. The biomarker was associated with tumor size, as follows. Patients in the large tumor size group were biomarker negative with probability 0.7 and biomarker positive with probability 0.3. Patients with small tumor size were biomarker negative with probability 0.3 and biomarker positive with probability 0.7. Let $Y \sim \text{LN}(m, s)$ denote a lognormal random variable $Y = \log T$ for $T \sim \text{N}(m, s)$. By a slight abuse of notation, we also use $\text{LN}(m, s)$ to denote the lognormal p.d.f. Let $\boldsymbol{x}_i = (1, x_{i,1}, x_{i,2}, x_{i,3})$ denote the covariates for patient i, here we include 1 in the covariate to indicate the intercept. We simulated each sample $Y_1, \cdots, Y_n$ of n observations from a mixture of lognormal distributions, $Y_i|\boldsymbol{x}_i \sim 0.4\ \text{LN}(\boldsymbol{x}_i\boldsymbol{\beta}_1, \sigma^2) + 0.6\ \text{LN}(\boldsymbol{x}_i\boldsymbol{\beta}_2, \sigma^2)$, where the true covariate parameters of the mixture components were $\boldsymbol{\beta}_1 = (1, 2, -2, 1)'$ and $\boldsymbol{\beta}_2 = (2, -1, 3, -3)'$, with $\sigma^2 = 0.4$. For comparison, we also fit an AFT regression model, assuming

$$Y_i \ = \ \log(T_i) = \boldsymbol{x}_i'\boldsymbol{\beta} + \sigma\epsilon_i, \quad i = 1, \ldots, n$$

with ϵ_i following an extreme value distribution, so that T_i follows a Weibull distribution.

In this simulation, we considered four scenarios, with $n = 50, 100$, or 200 observations without censoring or $n = 200$ with 23% censoring. For each scenario, $N = 1,000$ trials were simulated. For each simulated data set we fit a DDP-GP survival regression model $F(Y_i \mid \boldsymbol{x}_i)$. For simulation j, let $\overline{S}(t \mid \boldsymbol{x}) = p(T_{n+1} \geq t \mid \boldsymbol{x}_{n+1,j} = \boldsymbol{x}, data)$ denote the posterior expected survival function for a future patient with covariate $\boldsymbol{x}$. Using the empirical distribution $\frac{1}{n}\sum_{i=1}^{n} \delta_{\boldsymbol{x}_{ij}}$ to marginalize w.r.t. $\boldsymbol{x}_{n+1,j}$ and averaging across simulations, we get

$$\overline{S}(t) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{n}\sum_{i=1}^{n}\overline{S}(t \mid \boldsymbol{x}_{ij}).$$

Figure S1 compares $\overline{S}(\cdot)$ under the DDP-GP model with the simulation truth

$$S_0(t) = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{n}\sum_{i=1}^{n}S_0(t \mid \boldsymbol{x}_{ij}),$$

and maximum likelihood estimates (MLE) under Weibull AFT, Lognormal AFT, and Exponential AFT models. In each scenario, the true curve is given as a solid black solid line, the MLE of the survival functions under the AFT regression model assuming Weibull distribution, Lognormal distribution and Exponential distribution as green, blue, magenta solid lines respectively, and the posterior mean survival function under the DDP-GP model as a solid red line with point-wise 90% credible bands as two dotted red lines.

In all four scenarios, the DDP-GP model based estimate reliably recovered the shape of the true survival function and avoided the excessive bias seen with the Weibull, lognormal and exponential MLE. As expected, the three scenarios without censoring show that increasing sample size gives more accurate estimation. With 23% censoring, the DDP-GP estimate becomes less accurate, but it still is much closer to the simulation truth than the AFT regression models with Weibull, lognormal and exponential distributions.

D: Computing Mean Survival Time

The risk sets of the seven transition time in the leukemia trial are defined as follows. Let $G^0 = \{1, \ldots, n\}$ denote the initial risk set at the start of induction chemotherapy, and $G^{(0,r)} = \{i : \ s_{1i} = r\}$ for $r = D, C, R$, so $G^0 = G^{(0,D)} \cup G^{(0,C)} \cup G^{(0,R)}$. Similarly, $G^{(C,P)} =$

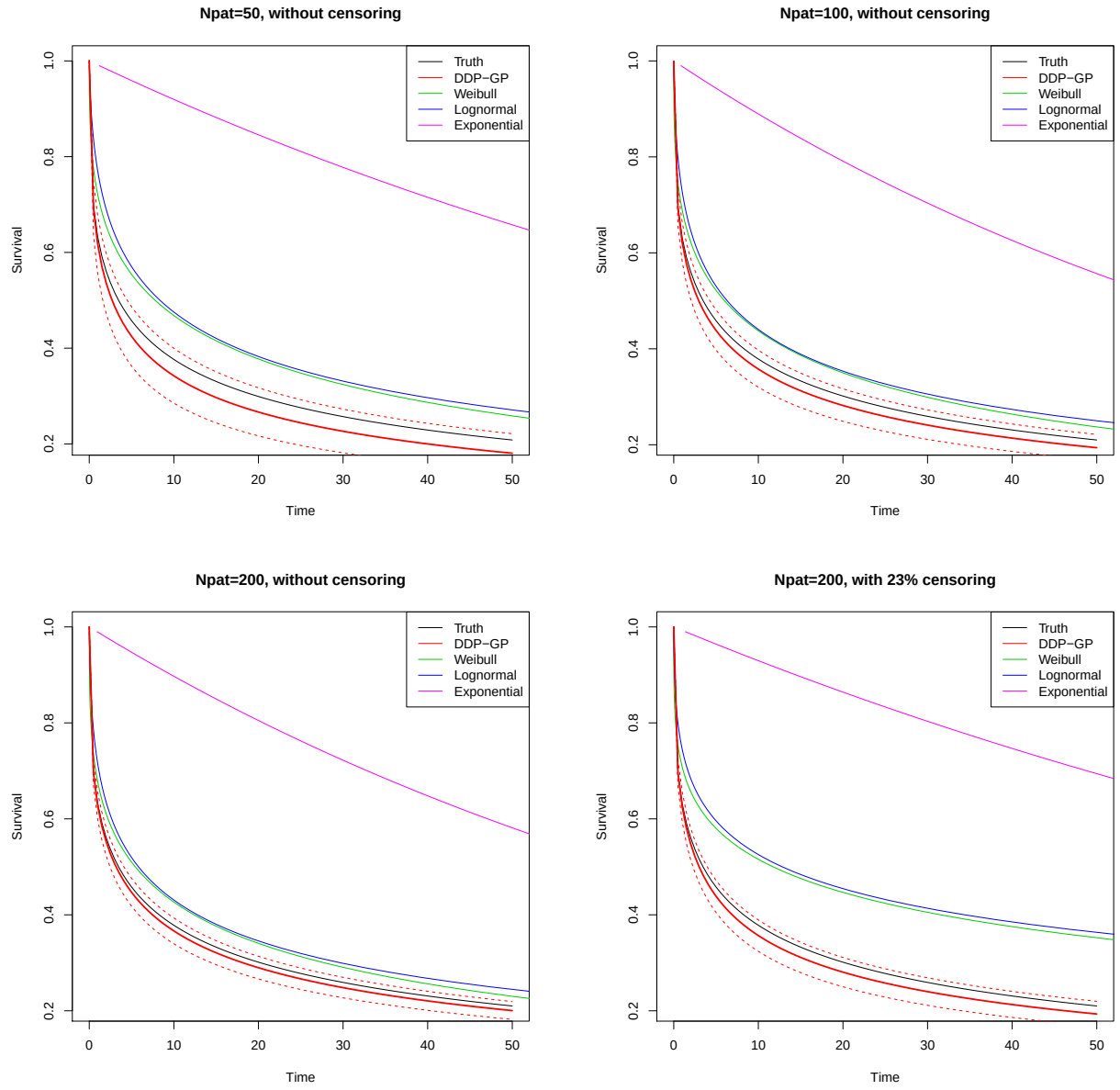

Figure S1: Simulation 1. True mean survival functions (black color) and estimated mean survival functions under the DDP-GP model (red color) for sample sizes $n = 50, 100, 200$ and $n = 200$ with 23% censoring for 1,000 simulations. For comparisons, we also show the MLE under an AFT regression with Weibull distribution (green color), Lognormal distribution (blue color) and Exponential distribution (magenta) . In all cases, the point-wise 90% credible bands are also displayed as the region between two dotted red lines.

$\{i : \ s_{1i} = C, s_{2i} = P\}$ is the later risk set for $T^{(P,D)}$.

To record right censoring, let U_i denote the time from the start of induction to last

followup for patient i. We assume that U_i is conditionally independent of the transition times given prior transition times and other covariates. Censoring of event times occurs by competing risk and/or loss to follow up. For a patient i in the risk set for event time T_i^k, let $\delta_i^k = 1$ if a patient i is not censored and 0 if patient i is right censored. For example, $\delta_i^{(0,D)} = 1$ for $i \in G^0$ if $T_i^{(0,D)} = \min(U_i, T_i^{(0,k)}, k = D, C, R)$. Similarly, $\delta_i^{(R,D)} = 1$ for $i \in G^{(0,R)}$ if $T_i^{(0,R)} + T_i^{(R,D)} < U_i$ and $\delta_i^{(P,D)} = 1$ for $i \in G^{(C,P)}$ if $T_i^{(0,C)} + T_i^{(C,P)} + T_i^{(P,D)} < U_i$.

Let $V_{x,i}$ denote the observed time for patient i in risk set G^x, as follows. For $i \in G^0$ let $V_{1,i} = \min(T_i^{(0,D)}, T_i^{(0,R)}, T_i^{(0,C)}, U_i)$ denote the observed time for the stage 1 event or censoring. For $i \in G^{(0,C)}$ let $V_{Ci} = \min(T_i^{(C,D)}, T_i^{(C,P)}, U_i - T_i^{(0,C)})$ denote the observed event time for the competing risks D and P and loss to followup. Similarly, for $i \in G^{(0,R)}$, let $V_{Ri} = \min(T_i^{(R,D)}, U_i - T_i^{(R,D)})$, and for $i \in G^{(C,P)}$ let $V_{(C,P),i} = \min(T_i^{(P,D)}, U_i - T_i^{(0,C)} - T_i^{(C,P)})$.

The joint likelihood function is the product $\mathcal{L} = \mathcal{L}_1\mathcal{L}_2\mathcal{L}_3\mathcal{L}_4$. The first factor $\mathcal{L}_1$ corresponds to response to induction therapy,

$$\mathcal{L}_1 = \prod_{i \in G^0} \prod_{k \in \{D,R,C\}} f^{(0,k)}(V_{1,i} \mid \boldsymbol{x}_i^{(0,k)})^{\delta_i^{(0,k)}} S^{(0,k)}(V_{1,i} \mid \boldsymbol{x}_i^{(0,k)})^{1-\delta_i^{(0,k)}}. \tag{5}$$

where $S^k = 1 - F^k$. The second factor $\mathcal{L}_2$ corresponds to patients $i \in G^{(0,R)}$ who experience resistance to induction and receive salvage $Z^{2,1}$,

$$\mathcal{L}_2 = \prod_{i \in G^{(0,R)}} f^{(R,D)}(V_{Ri} \mid \boldsymbol{x}_i^{(R,D)})^{\delta_i^{(R,D)}} S^{(R,D)}(V_{Ri} \mid \boldsymbol{x}_i^{(R,D)})^{1-\delta_i^{(R,D)}}. \tag{6}$$

The third factor $\mathcal{L}_3$ is the likelihood contribution from patients achieving CR,

$$\mathcal{L}_3 = \prod_{i \in G^{(0,C)}} \prod_{k=(C,D),(C,P)} f^k(V_{Ci} \mid \boldsymbol{x}_i^k)^{\delta_i^k} S^k(V_{Ci} \mid \boldsymbol{x}_i^k)^{1-\delta_i^k}. \tag{7}$$

The fourth factor $\mathcal{L}_4$ is the contribution from patients who experience tumor progression after CR

$$\mathcal{L}_4 = \prod_{i \in G^{(C,P)}} f^{(P,D)}(V_{CP,i} \mid \boldsymbol{x}_i^{(P,D)})^{\delta_i^{(P,D)}} S^{(P,D)}(V_{CP,i} \mid \boldsymbol{x}_i^{(P,D)})^{1-\delta_i^{(P,D)}}. \tag{8}$$

The mean survival time of a patient treated with regime $\mathbf{Z} = (Z^1, Z^{2,1}, Z^{2,2})$ is

$$\begin{aligned}\eta(\mathbf{Z}) = &\int \Big[p(s_1 = D \mid \boldsymbol{x}^0, Z^1)\eta^{(0,D)}(\boldsymbol{x}^0, Z^1)\Big] d\widehat{p}(\boldsymbol{x}^0) \\ &+ \int \Big\{p(s_1 = R \mid \boldsymbol{x}^0, Z^1)\Big[\eta^R(\boldsymbol{x}^0, Z^1) + \int \eta^{(R,D)}(\boldsymbol{x}^0, Z^1, Z^{2,1}, T^{(0,R)}) d\mu(T^{(0,R)})\Big]\Big\} d\widehat{p}(\boldsymbol{x}^0) \\ &+ \int p(s_1 = C \mid \boldsymbol{x}^0, Z^1)\Big[\eta^C(\boldsymbol{x}^0, Z^1) + \int \Big[p(s_2 = D \mid s_1 = C, \boldsymbol{x}^0, Z^1, T^{(0,C)})\eta^{(C,D)}(\boldsymbol{x}^0, Z^1, T^C) \\ &\qquad + p(s_2 = P \mid s_1 = C, \boldsymbol{x}^0, Z^1, T^{(0,C)})[\eta^{(C,P)}(\boldsymbol{x}^0, Z^1, T^{(0,C)}) \\ &\qquad\quad + \int \eta^{(P,D)}(\boldsymbol{x}^0, Z^1, Z^{2,2}T^{(0,C)}, T^{(C,P)}) d\mu(T^{(C,P)})]d\mu(T^{(0,C)})\Big] d\widehat{p}(\boldsymbol{x}^0). \quad (9)\end{aligned}$$

We compute the IPTW estimates for overall mean survival with regime $\mathbf{Z}$ as

$$IPTW(\mathbf{Z}) = \sum_{i=1}^{n} w_i(\mathbf{Z}) T_i \;/\; \sum_{i=1}^{n} w_i(\mathbf{Z}), \quad (10)$$

where

$$\begin{aligned} w_i(\mathbf{Z}) \;=\; & \frac{I(\mathbf{Z} = \mathbf{Z}_i)\delta_i}{\hat{K}(U_i)}\Big[I(s_{1i} = D) + I(s_{1i} = R) I_i(Z^{2,1}) / \hat{\Pr}(Z^{2,1} \mid s_{1i} = R, Z^1, \boldsymbol{x}_i^0, T_i^{(0,R)}) \\ & + I(s_{1i} = C, s_{2i} = D) \\ & + I(s_{1i} = C, s_{2i} = P) I_i(Z^{2,2}) / \hat{\Pr}(Z^{2,2} \mid s_{1i} = C, s_{2i} = P, Z^1, \boldsymbol{x}_i^0, T_i^{(0,C)}, T_i^{(C,P)})\Big]. \end{aligned} \quad (11)$$

In (11), $\hat{K}$ is the Kaplan-Meier estimator of the censoring survival distribution $K(u) = P(U \geq t)$ at time t. $I_i(Z)$ is is an indictor of treatment Z and 0 otherwise, and $\hat{\Pr}(Z^{2,1} \mid s_{1i} = C, Z^1, \boldsymbol{x}_i^0, T_i^{(0,R)})$ is the probability of receiving salvage treatment $Z^{2,1}$ estimated using logistic regression, and similarly for $\hat{\Pr}(Z^{2,2} \mid s_{1i} = C, s_{2i} = P, Z^1, \boldsymbol{x}_i^0, T_i^{(0,C)}, T_i^{(C,P)})$. The above estimator has been shown to be consistent under suitable assumptions (Wahed and Thall, 2013; Scharfstein et al., 1999).

E: Survival Regression for $T^{(C,P)}$ and $T^{(P,D)}$

Here we summarize results for the survival regression for $T^{(C,P)}$. Among the $n = 210$ patients, 102 (48.6%) achieved C, with C rates of 37%, 48%, 53% and 56% in the FAI, FAI plus ATRA, FAI plus GCSF and FAI plus GCSF plus ATRA arms, respectively. Of the 102 patients who achieved CR, 93 experienced disease progression before death or being lost to follow-up. Among these 93 relapsed patients, 53 received salvage treatment with HDAC. For a hypothetical future patient at age 61 and poor cytogenetic abnormality, Figure S2 summarizes survival regression functions for each of the four induction therapies, with solid lines representing $T^{(0,C)} = 20$ and dashed lines representing $T^{(0,C)} = 30$. The four dashed lines are below the four corresponding solid lines, indicating that $T^{(0,C)}$ was associated with $T^{(C,P)}$. This observation coincides with the well-known phenomenon in chemotherapy for AML or MDS that, regardless of induction therapy, the longer it takes to achieve C, the shorter the period that the patient remains in C.

Similarly, we summarize results for the survival regression for $T^{(P,D)}$. For a patient with poor cytogenetic abnormality, Figure S3 shows the posterior predicted survival functions under different combinations of induction therapy and age. Panels (a) and (c) show the survival functions of a patient assigned salvage treatment HDAC with age 46 or 76, while panels (b) and (d) plot the corresponding survival functions for the patient assigned non HDAC as salvage. Four different colors represent the four induction therapies. Figure S3 shows that residual time to D after disease progression following C was associated with both age and salvage therapy. Older patients are more likely to have shorter residual life once their disease progressed, and patients given HDAC as salvage die more quickly than patients given non HDAC salvage.

References

Ishwaran, H. and James, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association*, 96(453).

MacEachern, S. N. and Müller, P. (1998). Estimating mixture of dirichlet process models. *Journal of Computational and Graphical Statistics*, 7(2):223–238.

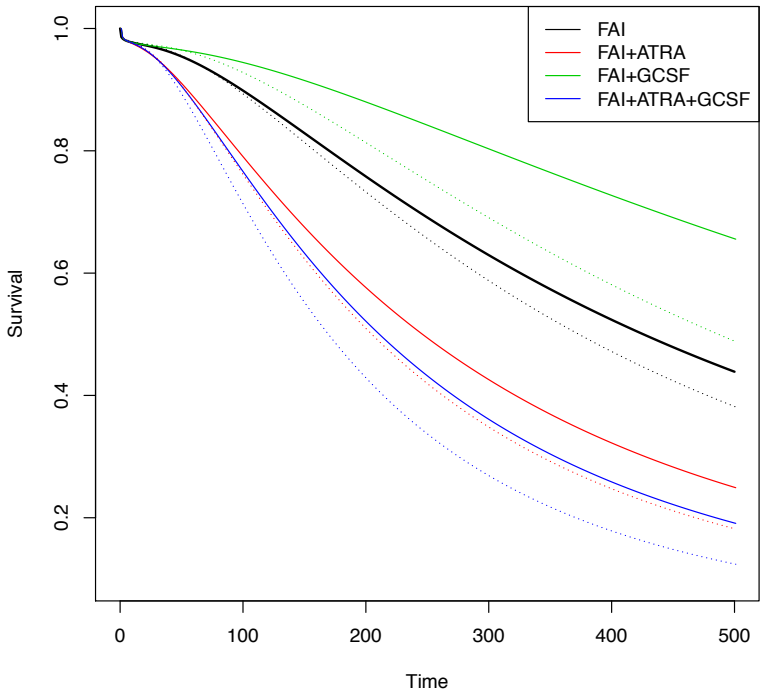

Figure S2: The effect of $T^{(0,C)}$ on $T^{(C,P)}$ at age 61 with poor cytogenetic abnormality. Black, red, green and blue curves represent induction treatments FAI, FAI+ATRA, FAI+GCSF and FAI+ATRA+GCSF, respectively. Solid lines and dash lines represent $T^{(0,C)} = 20$ and $T^{(0,C)} = 30$, respectively. The longer it takes to achieve C, the shorter the period of time that the patient remained in C.

Neal, R. M. (2000). Markov chain sampling methods for dirichlet process mixture models. *Journal of computational and graphical statistics*, 9(2):249–265.

Scharfstein, D. O., Rotnitzky, A., and Robins, J. M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models. *Journal of the American Statistical Association*, 94(448):1096–1120.

Wahed, A. S. and Thall, P. F. (2013). Evaluating joint effects of induction–salvage treatment regimes on overall survival in acute leukaemia. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 62(1):67–83.

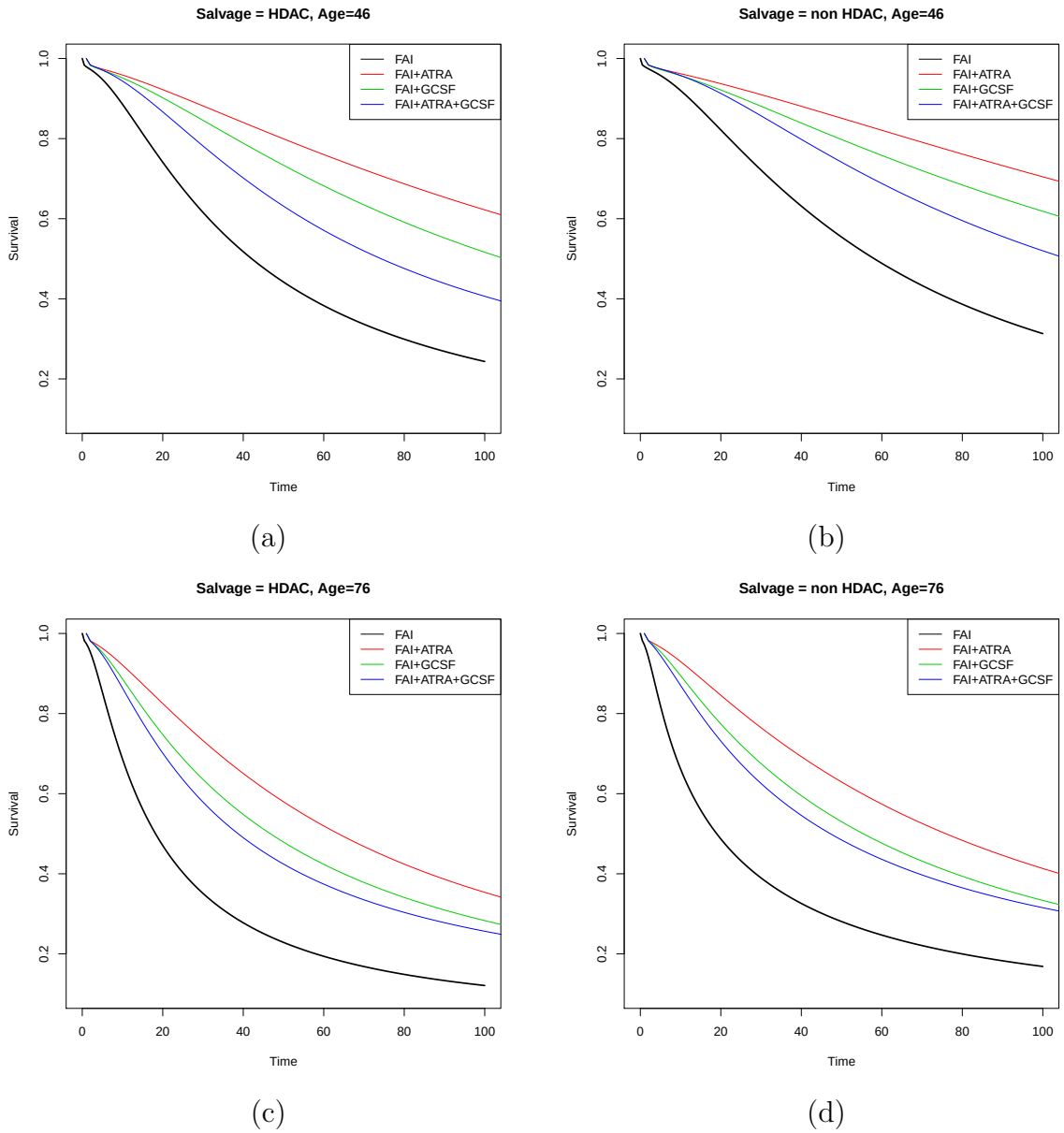

Figure S3: AML-MDS trial data in transition (P, D): Panels (a) and (c) show the posterior estimated survival functions of patient at age 46 and 76 with poor cytogenetic abnormality assigned to salvage treatment HDAC for four induction therapies respectively. Panels (b) and (d) show the posterior estimated survival functions of patient at age 46 and 76 with poor cytogenetic abnormality assigned to salvage treatment non HDAC for four induction therapies respectively. Black, red, green and blue curves represent induction treatments FAI, FAI+ATRA, FAI+GCSF and FAI+ATRA+GCSF, respectively.