Supplementary Web Materials for "A Bayesian Semiparametric Approach for Incorporating Longitudinal Information on Exposure History for Inference in Case-Control Studies"

1 Web Appendix

A.1: Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm for adaptive knot selection in the exposure trajectory and disease risk models

Recall the exposure trajectory model is given by

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1(t_{ij} + a_i^d) + \sum_{k=1}^{K} \beta_{k+1}(t_{ij} + a_i^d - \tau_k)_+ + b_i(t_{ij} + a_i^d) + e_{ij} \\ &= \mathbf{\Phi}(t_{ij} + a_i^d)'\boldsymbol{\beta} + b_i(t_{ij} + a_i^d) + e_{ij}, \end{aligned} \tag{1}$$

where $e_{ij} \sim N(0, \sigma_e^2)$ and $b_i \sim N(0, \sigma_b^2)$. Further, the disease risk model is given by

$$\begin{aligned} P(D_i = 1 | X_i(t + a_i^d), -c_1 \le t \le -c_2) &= L\left(\alpha + \delta a_i^d + \int_{-c_1}^{-c_2} X_i(t + a_i^d)\gamma(t)dt\right) \\ &= L\left(A_i^{d'}\boldsymbol{\theta} + \boldsymbol{\beta}' M_i \boldsymbol{\phi} + b_i Q_i \boldsymbol{\phi}\right), \end{aligned} \tag{2}$$

where $\boldsymbol{\theta} = (\alpha, \delta)'$, $A_i^d = (1, a_i^d)'$, $M_i = \int_{-c_1}^{-c_2} \mathbf{\Phi}(t + a_i^d)\mathbf{\Psi}(t)'dt$, $Q_i = \int_{-c_1}^{-c_2} (t + a_i^d)\mathbf{\Psi}(t)'dt$, and $\gamma(t) = \mathbf{\Psi}(t)'\phi$. Let $\boldsymbol{Y} = (Y_{i1}, Y_{i2}, \ldots, Y_{N,n_N})'$, $\boldsymbol{D} = (D_1, D_2, \ldots, D_N)'$, and $\boldsymbol{A}^d = (A_1^d, A_2^d, \ldots, A_N^d)'$. Also let $D_{obs} = (\boldsymbol{Y}, \boldsymbol{D}, \boldsymbol{A}^d)$ denote the observed data.

Likelihood Approximation

Following Albert and Chib (1993), we introduce latent variables $W_1, W_2, ..., W_N$ such that $D_i = 1$ if $W_i > 0$ and $D_i = 0$ otherwise. Let W_i be independently distributed from a t distribution with location $H_i = A_i^{d'}\boldsymbol{\theta} + \boldsymbol{\beta}' M_i \boldsymbol{\phi} + b_i Q_i \boldsymbol{\phi}$, scale parameter 1 and degrees of freedom ν. Equivalently, with the introduction of the additional random variable λ_i, the distribution of W_i can be expressed as scale mixtures of normal distributions, i.e.,

$$W_i | \lambda_i \sim N(H_i, \lambda_i^{-1}), \qquad \lambda_i \sim \text{Gamma}(\nu/2, 2/\nu)$$

where the Gamma pdf is proportional to $\lambda_i^{\nu/2-1}\exp(-\nu\lambda_i/2)$. Using this approximation, we can replace the logit link by the mixture of normals given above.

Case I : Sampling from the posterior conditional on the number and location of knots

Using the exposure trajectory model in (1), the disease risk model in (2) and using the above data augmentation algorithm, the likelihood function conditional on the known number and location of knots $(k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi})$ is given by

$$\begin{aligned} & L(\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2 | D_{obs}) \\ & = \prod_{i=1}^{N} \prod_{j=1}^{n_i} p(Y_{ij} | S_{ij}, \sigma_e^2) \times \prod_{i=1}^{N} \Delta_i \; p(W_i | H_i, 1/\lambda_i) \; G(\lambda_i | \nu/2, 2/\nu) \; p(b_i | 0, \sigma_b^2), \end{aligned} \tag{3}$$

where $\Delta_i = [I(W_i > 0) I(D_i = 1) + I(W_i \leq 0) I(D_i = 0)]$, $S_{ij} = \boldsymbol{\Phi}(t_{ij} + a_i^d)' \boldsymbol{\beta} + b_i (t_{ij} + a_i^d)$, and $H_i = A_i^{d'} \boldsymbol{\theta} + \boldsymbol{\beta}' M_i \boldsymbol{\phi} + b_i Q_i \boldsymbol{\phi}$. We note that $p(U|a, b)$ denotes a normal density with mean a and variance b and $G(V|a, b)$ denotes a gamma density with shape a and rate b. Let $\boldsymbol{W} = (W_1, W_2, \ldots, W_N)'$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)'$, and $\boldsymbol{b} = (b_1, b_2, \ldots, b_N)'$.

We assume that the joint prior for $(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2)$ is of the form $\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2) = \pi(\boldsymbol{\theta}) \pi(\boldsymbol{\beta} | \sigma_\beta^2) \pi(\sigma_\beta^2) \pi(\boldsymbol{\phi} | \sigma_\phi^2) \pi(\sigma_\phi^2) \pi(\sigma_e^2) \pi(\sigma_b^2)$. In particular,

$$\begin{aligned} & \boldsymbol{\theta} \sim N(\mathbf{0}, \sigma_\theta^2 I), \;\; \boldsymbol{\beta} \sim N(\mathbf{0}, \sigma_\beta^2 I), \;\; \boldsymbol{\phi} \sim N(\mathbf{0}, \sigma_\phi^2 I) \\ & \pi(\sigma_e^2) \propto (\sigma_e^2)^{-(a_0+1)} e^{-b_0/\sigma_e^2}, \;\; \pi(\sigma_b^2) \propto (\sigma_b^2)^{-(a_1+1)} e^{-b_1/\sigma_b^2}, \\ & \pi(\sigma_\beta^2) \propto (\sigma_\beta^2)^{-(a_2+1)} e^{-b_2/\sigma_\beta^2}, \;\; \text{and} \;\; \pi(\sigma_\phi^2) \propto (\sigma_\phi^2)^{-(a_3+1)} e^{-b_3/\sigma_\phi^2}, \end{aligned} \tag{4}$$

where σ_θ^2, a_0, b_0, a_1, b_1, a_2, b_2, a_3, and b_3 are the prespecified hyperparameters. In Section 6, we use $\sigma_\theta^2 = 100$ for $\pi(\boldsymbol{\theta})$, $a_0 = 0.1$ and $b_0 = 0.1$ for $\pi(\sigma_e^2)$, $a_1 = 0.1$ and $b_1 = 0.1$ for $\pi(\sigma_b^2)$, $a_2 = 3$ and $b_2 = 3$ for $\pi(\sigma_\beta^2)$, and $a_2 = 3$ and $b_2 = 3$ for $\pi(\sigma_\phi^2)$. Further, we also consider other values for (a_2, b_2) and (a_3, b_3) such as $(0.1, 0.1)$, $(1, 1)$, $(2, 2)$, and $(4, 4)$. Note that inferences are not very sensitive to the choice of hyperparameters.

Based on the joint prior distributions in (4), the joint posterior distribution of $\boldsymbol{W}$, $\boldsymbol{\lambda}$, $\boldsymbol{b}$, $\boldsymbol{\theta}$, $\boldsymbol{\beta}$, $\boldsymbol{\phi}$, σ_e^2, σ_b^2, σ_β^2, and σ_ϕ^2 based on the observed data D_{obs} is given by

$$\begin{aligned} \pi(\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2 | D_{obs}) & \propto L(\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2 | D_{obs}) \\ & \times \pi(\boldsymbol{\theta}) \pi(\boldsymbol{\beta} | \sigma_\beta^2) \pi(\sigma_\beta^2) \pi(\boldsymbol{\phi} | \sigma_\phi^2) \pi(\sigma_\phi^2) \pi(\sigma_e^2) \pi(\sigma_b^2), \end{aligned} \tag{5}$$

where $L(\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2 | D_{obs})$ is given in (3). Next, we develop an efficient Markov Chain Monte Carlo algorithm for fixed $(k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi})$.

To sample from the joint posterior $\pi(\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2 | D_{obs})$ given in (5), we require sampling from the conditional posterior distributions as follows: (i) $\pi(\boldsymbol{W} | \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2, D_{obs})$; (ii) $\pi(\boldsymbol{\lambda} | \boldsymbol{\theta}, \boldsymbol{b}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2, D_{obs})$; (iii) $\pi(\boldsymbol{b}, \sigma_b^2 | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_\beta^2, \sigma_\phi^2, D_{obs})$; (iv) $\pi(\sigma_e^2 | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2, D_{obs})$; (v) $\pi(\boldsymbol{\theta} | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2, D_{obs})$; (vi) $\pi(\boldsymbol{\beta},$

$\sigma^2_\beta|\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_b, \sigma^2_\phi, D_{obs})$; and (vii) $\pi(\boldsymbol{\phi}, \sigma^2_\phi|\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2_e, \sigma^2_b, \sigma^2_\beta, D_{obs})$. We briefly discuss how to sample from each of the conditional posterior distributions in (5). For (i),

$$W_i|\boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_b, \sigma^2_\beta, \sigma^2_\phi, D_{obs} \sim N\left(A_i^{d'}\boldsymbol{\theta} + \boldsymbol{\beta}' M_i \boldsymbol{\phi} + b_i Q_i \boldsymbol{\phi}, \frac{1}{\lambda_i}\right)\Delta_i.$$

Thus, sampling W_i from the conditional distributions $W_i|\boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_b, \sigma^2_\beta, \sigma^2_\phi, D_{obs}$ is straightforward. For (ii),

$$\lambda_i|\boldsymbol{\theta}, \boldsymbol{b}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_b, \sigma^2_\beta, \sigma^2_\phi, D_{obs} \sim \text{Gamma}\left(\frac{\nu+1}{2}, \frac{\nu + (W_i - A_i^{d'}\boldsymbol{\theta} - \boldsymbol{\beta}' M_i \boldsymbol{\phi} - \boldsymbol{b}_i' Q_i \boldsymbol{\phi})}{2}\right).$$

For (iii), we apply the collapsed Gibbs technique of Liu (1994) via the following identity:

$$\begin{aligned}\pi(\boldsymbol{b}, \sigma^2_b|\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_\beta, \sigma^2_\phi, D_{obs}) &= \pi(\boldsymbol{b}|\sigma^2_b, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_\beta, \sigma^2_\phi, D_{obs}) \\ &\quad \times \pi(\sigma^2_b|\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_\beta, \sigma^2_\phi, D_{obs}).\end{aligned}$$

That is, we sample σ^2_b after collapsing out $\boldsymbol{b}$. Given $\sigma^2_b, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_\beta, \sigma^2_\phi$, and D_{obs}, b_i's are conditionally independent and

$$b_i|\sigma^2_b, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_\beta, \sigma^2_\phi, D_{obs} \sim N(A_b^{-1}B_b, A_b^{-1}),$$

where

$$\begin{aligned}A_b &= \frac{1}{\sigma^2_e}\sum_j (t_{ij} + a_i^d)^2 + \frac{1}{\sigma^2_b} + \lambda_i \boldsymbol{\phi} Q_i' Q_i \boldsymbol{\phi} \quad \text{and} \\ B_b &= \frac{1}{\sigma^2_e}\sum_j (y_{ij} - \boldsymbol{\Phi}(t_{ij} + a_i^d)'\boldsymbol{\beta})(t_{ij} + a_i^d) + \lambda_i (W_i - A_i^{d'}\boldsymbol{\theta} - \boldsymbol{\beta}' M_i \boldsymbol{\phi}) Q_i \boldsymbol{\phi}.\end{aligned}$$

Thus, sampling b_i from the conditional distributions $\pi(b_i|\sigma^2_b, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_\beta, \sigma^2_\phi, D_{obs})$ is straightforward. Further, the conditional posterior density for $\pi(\sigma^2_b|\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_\beta, \sigma^2_\phi, D_{obs})$ has the form

$$\begin{aligned}&\pi(\sigma^2_b|\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_\beta, \sigma^2_\phi, D_{obs}) \\ &\propto (\sigma^2_b)^{-(N/2 + a_1 + 1)} \exp\left(-\frac{b_1}{\sigma^2_b}\right) \times \prod_i A_i^{-1/2} \exp\left(\frac{B_b^2}{2A_b}\right).\end{aligned}$$

We use the Metropolis-Hastings algorithm to sample σ^2_b from conditional distribution $\pi(\sigma^2_b|\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2_e, \sigma^2_\beta, \sigma^2_\phi, D_{obs})$. For (iv),

$$\sigma^2_e \sim \text{IG}\left(a_0 + \sum_i \sum_j \frac{1}{2} + 1, b_0 + \frac{1}{2}\sum_i \sum_j \left(y_{ij} - \boldsymbol{\Phi}(t_{ij} + a_i^d)'\boldsymbol{\beta} - b_i(t_{ij} + a_i^d)\right)^2\right),$$

where IG denote an inverse-gamma distribution. For (vi), we also apply the collapsed Gibbs technique of Liu (1994) via the following identity:

$$\pi(\boldsymbol{\beta}, \sigma_\beta^2 | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\phi^2, D_{obs}) = \pi(\boldsymbol{\beta} | \sigma_\beta^2, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\phi^2, D_{obs}) \times \pi(\sigma_\beta^2 | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\phi^2, D_{obs}).$$

That is, we sample σ_β^2 after collapsing out $\boldsymbol{\beta}$. Given $\sigma_\beta^2, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\phi^2$, and D_{obs}, the conditional distribution of $\boldsymbol{\beta}$ is the normal distribution as

$$\boldsymbol{\beta} | \sigma_\beta^2, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\phi^2, D_{obs} \sim N(A_E^{-1} B_E, A_E^{-1}),$$

where

$$A_E = \frac{1}{\sigma_e^2} \sum_i^N \sum_j^{n_i} \boldsymbol{\Phi}(t_{ij} + a_i^d) \boldsymbol{\Phi}(t_{ij} + a_i^d)' + \sum_i^N \lambda_i M_i \boldsymbol{\phi} \boldsymbol{\phi}' M_i' + \frac{1}{\sigma_\beta^2} I,$$

$$B_E = \frac{1}{\sigma_e^2} \sum_i^N \sum_j^{n_i} \boldsymbol{\Phi}(t_{ij} + a_i^d)(y_{ij} - b_i(t_{ij} + a_i^d)) + \sum_i^N \lambda_i M_i \boldsymbol{\phi}(W_i - A_{di}' \boldsymbol{\theta} - b_i Q_i \boldsymbol{\phi}).$$

Further, the conditional posterior density for $\pi(\sigma_\beta^2 | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\phi^2, D_{obs})$ has the form

$$\pi(\sigma_\beta^2 | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\phi^2, D_{obs}) \propto \left(\sigma_\beta^2\right)^{-(a_2 + \frac{p}{2} + 1)} \exp\left(-\frac{b_2}{\sigma_\beta^2}\right) \times |A_E|^{-1/2} \exp\left[\frac{1}{2} B_E' A_E B_E\right].$$

Again, we use the Metropolis-Hastings algorithm to sample σ_β^2 from conditional distribution $\pi(\sigma_\beta^2 | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\phi^2, D_{obs})$. For (vii), we also apply the collapsed Gibbs technique of Liu (1994) via the following identity:

$$\pi(\boldsymbol{\phi}, \sigma_\phi^2 | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, D_{obs}) = \pi(\boldsymbol{\phi} | \sigma_\phi^2, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, D_{obs}) \times \pi(\sigma_\phi^2 | \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, D_{obs}).$$

That is, we sample σ_ϕ^2 after collapsing out $\boldsymbol{\phi}$. Given $\sigma_\phi^2, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2$, and D_{obs}, the conditional distribution of $\boldsymbol{\phi}$ is the normal distribution as

$$\boldsymbol{\phi} | \sigma_\phi^2, \boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, D_{obs} \sim N(A_D^{-1} B_D, A_D^{-1}),$$

where

$$A_D = \sum_i^N \lambda_i (\boldsymbol{\beta}' M_i + b_i Q_i)' (\boldsymbol{\beta}' M_i + b_i Q_i) + \frac{1}{\sigma_\phi^2} I, \quad \text{and}$$

$$B_D = \sum_i^N \lambda_i (\boldsymbol{\beta}' M_i + b_i Q_i)' (W_i - A_{di}' \boldsymbol{\theta}).$$

Finally, the conditional posterior density for $\pi(\sigma_\phi^2|\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, D_{obs})$ has the form

$$\begin{aligned}
&\pi(\sigma_\phi^2|\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, D_{obs}) \\
&\propto \left(\sigma_\phi^2\right)^{-(a_3+\frac{q}{2}+1)} \exp\left(-\frac{b_3}{\sigma_\phi^2}\right) \times |A_D|^{-1/2} \exp\left[\frac{1}{2}B_D' A_D B_D\right].
\end{aligned}$$

which we sample using the Metropolis-Hastings algorithm.

Case II: Sampling from the posterior when the number and locations of knots are unknown.

Recall that k_1 and k_2 are the number of knots for the exposure and disease risk models, respectively, where $0 \le k_1 \le K_1$, $0 \le k_2 \le K_2$ with K_1 and K_2 prespecified constants. Let $\boldsymbol{\tau} = (\tau_1, ..., \tau_{k_1})'$ and $\boldsymbol{\xi} = (\xi_1, ..., \xi_{k_2})'$ denote the corresponding knot locations such that

$$a^E < \tau_1 < ... < \tau_{k_1} < b^E \;\text{ and }\; a^D < \xi_1 < ... < \xi_{k_2} < b^D.$$

We assume that the joint prior for $(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2, k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi})$ is of the form

$$\begin{aligned}
&\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2, k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi}) \\
&= \pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2|k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi})\pi(k_1)\pi(\boldsymbol{\tau}|k_1)\pi(k_2)\pi(\boldsymbol{\xi}|k_2), \qquad (6)
\end{aligned}$$

where $\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2|k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi})$ is given in (4). For our analysis, given k_1 and k_2, we assume a uniform prior distribution for $\boldsymbol{\tau}$ and $\boldsymbol{\xi}$, respectively as

$$\begin{aligned}
\pi(\boldsymbol{\tau}|k_1) &= \frac{k_1!}{(b^E - a^E)^{k_1}} I(a^E < \tau_1 < ... < \tau_{k_1} < b^E) \;\text{ and} \\
\pi(\boldsymbol{\xi}|k_2) &= \frac{k_2!}{(b^D - a^D)^{k_2}} I(a^D < \xi_1 < ... < \xi_{k_2} < b^D).
\end{aligned}$$

We assume the Poisson prior distribution with mean μ_1 and mean μ_2 for k_1 and k_2, respectively. For the analysis in Section 6, we used $\mu_1 = 1$ and $K_1 = 5$ for k_1, and $\mu_2 = 1$ and $K_2 = 5$ for k_2; inferences were not very sensitive to these choices. Based on the joint prior distributions in (6), the joint posterior distribution of $\boldsymbol{W}$, $\boldsymbol{\lambda}$, $\boldsymbol{b}$, $\boldsymbol{\theta}$, $\boldsymbol{\beta}$, $\boldsymbol{\phi}$, σ_e^2, σ_b^2, σ_β^2, σ_ϕ^2, k_1, $\boldsymbol{\tau}$, k_2, and $\boldsymbol{\xi}$ based on the observed data D_{obs} is thus given by

$$\begin{aligned}
&\pi(\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2, k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi}|D_{obs}) \\
&\propto L(\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi}|D_{obs}) \\
&\quad \times \pi(\boldsymbol{\theta})\pi(\boldsymbol{\beta}|\sigma_\beta^2)\pi(\sigma_\beta^2)\pi(\boldsymbol{\phi}|\sigma_\phi^2)\pi(\sigma_\phi^2)\pi(\sigma_e^2)\pi(\sigma_b^2), \pi(k_1)\pi(\boldsymbol{\tau}|k_1)\pi(k_2), \qquad (7)
\end{aligned}$$

where $L(\boldsymbol{W}, \boldsymbol{\lambda}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi}|D_{obs})$ is given in (3). Next, we develop an efficient reversible jump method to sample $(k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi})$.

Reversible Jump Markov Chain Monte Carlo

In the following, we develop an efficient Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm to deal with the changing dimension of the parameter space when adding or deleting knots. Our algorithm modifies and extends previous work (DiMateo et al., 2001; Botts and Daniels, 2008). The RJMCMC algorithm consists of three types of transitions: knot addition (birth step), knot deletion (death step), and knot relocation (relocation step). The probabilities for choosing each type of the moves are denoted b_k, d_k and ζ_k, respectively, and are given by

$$b_k = c \min\left\{1, \frac{\pi(k+1)}{\pi(k)}\right\}, \quad d_k = c \min\left\{1, \frac{\pi(k-1)}{\pi(k)}\right\}, \quad \text{and} \quad \zeta_k = 1 - b_k - d_k. \tag{8}$$

In the analysis, we take $c = 0.4$. For each step, the acceptance probability is given by $\alpha = \min\{1, \text{likelihood ratio} \times \text{prior ratio} \times \text{proposal ratio}\}$. To determine whether or not to move from state $(k_1, \boldsymbol{\tau}, k_2, \boldsymbol{\xi})$ to another state $(k_1^*, \boldsymbol{\tau}^*, k_2^*, \boldsymbol{\xi}^*)$ using the reversible jump MCMC method, we need to obtain the conditional posterior distributions of $(k_1, \boldsymbol{\tau})$ and $(k_2, \boldsymbol{\xi})$ after integrating out $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$ from joint posterior distribution in (7), respectively. The conditional posteriors of $(k_1, \boldsymbol{\tau})$ and $(k_2, \boldsymbol{\xi})$ are given by

$$\begin{aligned} &\pi(k_1, \tau | W, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2, D_{obs}) \propto |A_E|^{-1/2} \exp\left[\frac{1}{2} B_E' A_E^{-1} B_E\right] \quad \text{and} \\ &\pi(k_2, \xi | W, \boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \sigma_\beta^2, \sigma_\phi^2, D_{obs}) \propto |A_D|^{-1/2} \exp\left[\frac{1}{2} B_D' A_D^{-1} B_D\right]. \end{aligned} \tag{9}$$

where A_E, B_E, A_D, and B_D are given in Case I. The reversible jump Markov chain Monte Carlo algorithm operates as follows.

Step 1. Initialization: set $(k_1^{(0)}, \boldsymbol{\tau}^{(0)}, k_2^{(0)}, \boldsymbol{\xi}^{(0)}, \boldsymbol{\beta}^{(0)}, \boldsymbol{\phi}^{(0)})$ and $i = 0$.

Step 2. Iteration i: generate $u^* \sim u(0,1)$, choose either step 2.1 for $(k_1, \boldsymbol{\tau}, \boldsymbol{\beta})$ or step 2.2 for $(k_2, \boldsymbol{\xi}, \boldsymbol{\phi})$, and update the rest of the parameters in step 2.3. (Here $u(a,b)$, denotes a uniform distribution over (a, b)).

Step 2.1. If $u^* < p$, sample $(k_1, \boldsymbol{\tau}, \boldsymbol{\beta})$. To do this, generate $u^{**} \sim u(0,1)$ and do

- the birth step if $u^{**} < b_k$.
- the death step if $b_k \leq u^{**} < b_k + d_k$.
- the relocation step if $u^{**} \geq b_k + d_k$.
- then generate $\boldsymbol{\beta}$ conditional on $(k_1, \boldsymbol{\tau})$.

Step 2.2. If $u^* \geq p$, sample $(k_2, \boldsymbol{\xi}, \boldsymbol{\phi})$. To do this, generate $u^{**} \sim u(0,1)$ and do

- the birth step if $u^{**} < b_k$.
- the death step if $b_k \leq u^{**} < b_k + d_k$.
- the relocation step if $u^{**} \geq b_k + d_k$.
- then generate $\boldsymbol{\phi}$ conditional on $(k_2, \boldsymbol{\xi})$.

Step 2.3. Update the rest of the parameters as outlined in Case I.

Step 3 Let $i \leftarrow i + 1$ and go to step 2.

We now provide more details on steps 2.1 and 2.2. We set $p = 0.5$. Given a new set $(k_1, \boldsymbol{\tau})$ and $(k_2, \boldsymbol{\xi})$, we generate $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$ from the their conditional posterior distributions, respectively as $\boldsymbol{\beta} \sim N(A_E^{-1} B_E, A_E^{-1})$ and $\boldsymbol{\phi} \sim N(A_D^{-1} B_D, A_D^{-1})$. We use the following proposal scheme for the birth, death, and relocation steps. First we consider Step 2.1 for $(k_1, \boldsymbol{\tau})$. Let $M_\tau^{k_1} = \{k_1, \tau_1, ..., \tau_{k_1}\}$ denote the current model defined by k_1 and $\boldsymbol{\tau}$. Let ϵ denote a tuning constant which restricts knots to not be added too close to the current knot locations. As such, intervals for adding a knot are defined as $I_k = (\tau_{k-1} + \epsilon, \tau_k - \epsilon)$, $k = 1, \ldots, k_1 + 1$.

For the birth step, we choose the candidate interval I_k uniformly using existing knots. The new τ^* is generated from $\tau^* \sim u(\tau_{k-1} + \epsilon, \tau_k - \epsilon)$. Here, the jump probability is given by

$$q(M_\tau^{k_1+1} | M_\tau^{k_1}) = \frac{b_{k_1}}{\tau_k - \tau_{k-1} - 2\epsilon}.$$

For the death step, the deleted knot is chosen uniformly from the existing knots and the jump probability is

$$q(M_\tau^{k_1+1} | M_\tau^{k_1}) = \frac{d_{k_1}}{k_1}.$$

For the relocation step, we choose a knot τ_s uniformly from existing knots. We also choose a candidate interval I_k uniformly using the existing knots. Then a new τ_s^* is generated from $\tau_s^* \sim u(\tau_{k-1} + \epsilon, \tau_k - \epsilon)$. Here, the jump probability is given by

$$q(M_{\tau^*}^{k_1} | M_\tau^{k_1}) = \frac{\zeta_{k_1}}{\tau_k - \tau_{k-1} - 2\epsilon}.$$

Step 2.2 for (k_2, ϕ) is similar to Step 2.1. In analysis in Section 6, we choose $\epsilon = 5$ for $\boldsymbol{\tau}$ and $\epsilon = 1$ for $\boldsymbol{\phi}$.

A.2: Bayesian Equivalence of prospective and retrospective likelihoods in the semiparametric context:

As mentioned in Section 1.1, for certain choices of the priors on the log odds, posterior inference for the parameter of interest based on a prospective logistic model can be shown to be equivalent to that based on a retrospective one. As a result, a prospective modeling framework can be used to analyze case-control data which are generally collected retrospectively. Here we show that the Bayesian equivalence results of Seaman and Richardson (2004) can be extended to the semiparametric framework we have proposed. Infact, we show equivalence under a general setting which allows for other time constant and time dependent covariates (in addition to the exposure profile). This enables us to use a prospective logistic framework (as described in Section 3) to analyze the PSA dataset in Section 5.

Our modeling framework hinges on the idea that for every subject, instead of a single exposure observation, a series of past exposure observations are available. We use this "exposure trajectory" or "exposure profile" in analyzing the present disease status of a subject. In the spirit of our dataset, we assume that the exposure observations are continuous. Let the exposure profile for the i^{th} subject be $X_i(t) = \{X_{i1}(t), ..., X_{in_i}(t), i = 1, ..., N\}$ $(-c_1 \leq t \leq -c_2)$ where $X_{ij}(t)$ is the j^{th} exposure observation recorded for the i^{th} subject as a function of t. Since an exposure trajectory is composed of a finite set of exposure observations, the discretizing mechanism proposed by Rubin (1981) and later by Gustafson (2002) can be applied to the trajectory as a whole i.e $\{X_i(t), -c_1 \leq t \leq -c_2\}$ can be assumed to be a discrete random variable with support $\{Z_1(t), ..., Z_J(t), -c_1 \leq t \leq -c_2\}$, the set of all observable exposure trajectories where $\{Z_j(t), -c_1 \leq t \leq -c_2, j = 1, ..., J\}$ is a finite collection of elements in the support of the $X_{ij}(t)$'s. Let Y_{0j} and Y_{1j} be the number of controls and cases having exposure profile $\{Z_j(t), -c_1 \leq t \leq -c_2\}$. We denote the "Null" or "baseline" trajectory as $\{X(t) = 0, -c_1 \leq t \leq -c_2\}$.

We also assume the presence of other time constant and time varying covariates which may influence the current disease status. Thus, the general form of the disease model is given by

$$P(D = 1|X(t), U(t), -c_1 \leq t \leq -c_2, \boldsymbol{w}) = L\left(\alpha + \boldsymbol{\delta}'\boldsymbol{w} + \int_{-c_1}^{-c_2} U(t)\lambda(t)dt + \int_{-c_1}^{-c_2} X(t)\gamma(t)dt\right)$$

where $\boldsymbol{w}$ is a vector of time constant covariates (like age at diagnosis in equation (3) for instance) while $\{U(t), -c_1 \leq t \leq -c_2\}$ may be some other time varying covariate (distinct from the exposure profile) with $\lambda(t)$ expressing its (possibly) time varying influence pattern on the current disease state.

Thus, the odds of disease corresponding to $\{Z_j(t), -c_1 \leq t \leq -c_2\}$ is

$$\exp\left(\alpha + \boldsymbol{\delta}'\boldsymbol{w} + \int_{-c_1}^{-c_2} U(t)\lambda(t)dt + \int_{-c_1}^{-c_2} Z_j(t)\gamma(t)dt\right)$$

and that corresponding to baseline exposure is

$$\exp\left(\alpha + \boldsymbol{\delta}'\boldsymbol{w} + \int_{-c_1}^{-c_2} U(t)\lambda(t)dt\right)$$

So, the odds ratio of disease corresponding to $\{Z_j(t), -c_1 \leq t \leq -c_2\}$ with respect to baseline exposure is $\exp\left(\int_{-c_1}^{-c_2} Z_j(t)\gamma(t)dt\right)$. Assuming that a control has exposure profile $\{Z_j(t), -c_1 \leq t \leq -c_2\}$ and other extrinsic covariates $(\{U(t) = u(t), -c_1 \leq t \leq -c_2\}, \boldsymbol{w})$ with probability $\delta_j / \sum_{k=1}^{J} \delta_k$, it can be easily shown that

$$P(X(t) = Z_j(t), U(t) = u(t), -c_1 \leq t \leq -c_2, \boldsymbol{w}|D = 1) = \frac{\delta_j \exp\left(\int_{-c_1}^{-c_2} Z_j(t)\gamma(t)dt\right)}{\sum_{k=1}^{J} \delta_k \exp\left(\int_{-c_1}^{-c_2} Z_k(t)\gamma(t)dt\right)}$$

Thus, the retrospective likelihood is

$$L(\boldsymbol{\delta}, \boldsymbol{\phi}) = \prod_{d=0}^{1}\prod_{j=1}^{J}\left[\frac{\delta_j \exp\left(d\int_{-c_1}^{-c_2} Z_j(t)\gamma(t)dt\right)}{\sum_{k=1}^{J} \delta_k \exp\left(d\int_{-c_1}^{-c_2} Z_k(t)\gamma(t)dt\right)}\right]^{y_{dj}} = \prod_{d=0}^{1}\prod_{j=1}^{J}\left[\frac{\delta_j \exp\left(d\boldsymbol{\phi}'\int_{-c_1}^{-c_2} Z_j(t)\boldsymbol{\Psi}(t)dt\right)}{\sum_{k=1}^{J} \delta_k \exp\left(d\boldsymbol{\phi}'\int_{-c_1}^{-c_2} Z_k(t)\boldsymbol{\Psi}(t)dt\right)}\right]^{y_{dj}}$$

since $\gamma(t) = \boldsymbol{\Psi}(t)'\boldsymbol{\phi} = \boldsymbol{\phi}'\boldsymbol{\Psi}(t)$ by (4). We assume $\delta_1 = 1$ for identifiability. Here $d = 0$ and 1 stands for controls and cases respectively. Assuming ϑ to be the baseline odds of disease i.e

$$\begin{aligned}\vartheta &= \frac{P(D = 1|X(t) = 0, U(t) = u(t), -c_1 \leq t \leq -c_2, \boldsymbol{w})}{P(D = 0|X(t) = 0, U(t) = u(t), -c_1 \leq t \leq -c_2, \boldsymbol{w})} \\ &= \exp\left(\alpha + \boldsymbol{\delta}'\boldsymbol{w} + \int_{-c_1}^{-c_2} U(t)\lambda(t)dt\right)\end{aligned}$$

the prospective likelihood is given by

$$L(\vartheta, \boldsymbol{\phi}) = \prod_{d=0}^{1}\prod_{j=1}^{J}\left[\frac{\vartheta^d \exp\left(d\boldsymbol{\phi}'\int_{-c_1}^{-c_2} Z_j(t)\boldsymbol{\Psi}(t)dt\right)}{\sum_{k=0}^{1} \vartheta^k \exp\left(k\boldsymbol{\phi}'\int_{-c_1}^{-c_2} Z_k(t)\boldsymbol{\Psi}(t)dt\right)}\right]^{y_{dj}}$$

Based on the above setup, we have the following equivalence results :

Theorem 1 : *The profile likelihood of $\boldsymbol{\phi}$ obtained by maximizing $L(\boldsymbol{\delta}, \boldsymbol{\phi})$ with respect to $\boldsymbol{\delta}$ is the same as that obtained by maximizing $L(\vartheta, \boldsymbol{\phi})$ with respect to ϑ.*

Proof : Let Y_{dj} $(d = 0, 1; j = 1, ..., J)$ be independently distributed as Poisson(λ_{dj}) where

$$\log\lambda_{dj} = \log\mu + d\log\vartheta + \log\delta_j + d\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt \tag{10}$$

Thus, the likelihood will be

$$L(\mu, \vartheta, \boldsymbol{\delta}, \boldsymbol{\phi}) = \prod_{d=0}^{1}\prod_{j=1}^{J}(\lambda_{dj})^{y_{dj}}\exp(-\lambda_{dj})$$

and hence the log likelihood will be

$$l(\mu, \vartheta, \boldsymbol{\delta}, \boldsymbol{\phi}) = \sum_{d=0}^{1}\sum_{j=1}^{J}\{y_{dj}\log(\lambda_{dj}) - \lambda_{dj}\}$$

Now, replacing the expression of $\log\lambda_{dj}$ from (10) we have

$$\begin{aligned} l(\mu, \vartheta, \boldsymbol{\delta}, \boldsymbol{\phi}) &= \sum_{d=0}^{1}\sum_{j=1}^{J} y_{dj}\left(\log\mu + d\log\vartheta + \log\delta_j + d\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right) \\ &- \sum_{d=0}^{1}\sum_{j=1}^{J} \mu\vartheta^d\delta_j\exp\left(d\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right) \end{aligned} \tag{11}$$

Differentiating (11) w.r.t μ and ϑ and solving the resulting equations we have

$$\hat{\mu} = \textstyle\sum_j y_{0j}/\sum_j \delta_j \quad \text{and} \quad \hat{\vartheta} = \dfrac{\displaystyle\sum_j y_{1j}\sum_j \delta_j}{\displaystyle\sum_j y_{0j}\sum_j \delta_j\exp\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right)}$$

Replacing the above expressions in (11) and then exponentiating, we obtain the expression of $L(\boldsymbol{\delta}, \boldsymbol{\phi})$.

Again, differentiating (11) w.r.t δ_j, we have

$$\delta_j = \frac{\displaystyle\sum_d y_{dj}}{\displaystyle\mu\sum_d \vartheta^d\exp\left(d\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right)}, \quad j = 1, ..., J \tag{12}$$

It is easy to show that if we replace (12) in (11) and then exponentiate, we get the expression for $L(\vartheta, \boldsymbol{\phi})$. Since the order of maximization is immaterial, it follows that, $L(\boldsymbol{\delta}, \boldsymbol{\phi})$ and $L(\vartheta, \boldsymbol{\phi})$, once maximized over the nuisance parameters (ϑ and $\boldsymbol{\delta}$ respectively) yield the same profile likelihood of $\boldsymbol{\phi}$. Thus, inferences about the parameter of interest $\boldsymbol{\phi}$ can be obtained using the prospective likelihood which has fewer nuisance parameters than the retrospective one.

Theorem 2 : *Let Y_{dj} $(d = 0, 1; j = 1, ..., J)$ be independently distributed as Poisson(λ_{dj}) where*

$$\log\lambda_{dj} = d\log\vartheta + \log\delta_j + d\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\boldsymbol{\Psi}(t)dt \tag{13}$$

We assume independent priors, $p(\vartheta) \propto \vartheta^{-1}$ and $p(\delta_j) \propto \delta_j^{a_j-1}$ for ϑ and $\boldsymbol{\delta}$. The prior for $\boldsymbol{\phi}$, $p(\boldsymbol{\phi})$ is chosen to be independent of ϑ and $\boldsymbol{\delta}$ such that for some q and r such that $y_{0q} \geq 1$ and $y_{0r} \geq 1$, $E\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_q(t)\boldsymbol{\Psi}(t)dt\right)$ and $E\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_r(t)\boldsymbol{\Psi}(t)dt\right)$ exists and are finite (i.e $p(\boldsymbol{\phi})$ is such that $E(\boldsymbol{\phi})$ exists and is finite). Let $y_{+j} = y_{0j} + y_{1j}$ and $y_{d+} = \sum_{j=1}^{J} y_{dj}$. Then the following two statements hold :

(i) *Assuming $w = log\vartheta$, the posterior density of $(w, \boldsymbol{\phi})$ is*

$$p(w, \boldsymbol{\phi}|\boldsymbol{y}) \propto p(\boldsymbol{\phi}) \prod_{j=1}^{J} \frac{\left\{\exp\left(w + \boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\boldsymbol{\Psi}(t)dt\right)\right\}^{y_{1j}}}{\left\{1 + \exp\left(w + \boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\boldsymbol{\Psi}(t)dt\right)\right\}^{y_{+j}+a_j}} \tag{14}$$

(ii) *Assuming $\boldsymbol{\theta} = (\theta_1, ..., \theta_J)$ and $\theta_j = \delta_j / \sum_{k=1}^{J} \delta_k$, the posterior density of $(\boldsymbol{\theta}, \boldsymbol{\phi})$ is*

$$p(\boldsymbol{\theta}, \boldsymbol{\phi}|\boldsymbol{y}) \propto p(\boldsymbol{\phi}) \prod_{j=1}^{J} \theta_j^{a_j-1} \prod_{d=0}^{1} \left[\frac{\prod_{j=1}^{J} \left\{\theta_j \exp\left(d\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\boldsymbol{\Psi}(t)dt\right)\right\}^{y_{dj}}}{\left\{\sum_{j=1}^{J} \theta_j \exp\left(d\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\boldsymbol{\Psi}(t)dt\right)\right\}^{y_{d+}}} \right] \tag{15}$$

(iii) *The marginal posterior densities of $\boldsymbol{\phi}$ obtainable from $p(w, \boldsymbol{\phi}|\boldsymbol{y})$ and $p(\boldsymbol{\theta}, \boldsymbol{\phi}|\boldsymbol{y})$ are the same.*

Proof : (i) The posterior density of $(\vartheta, \boldsymbol{\delta}, \boldsymbol{\phi})$ is

$$p(\vartheta, \boldsymbol{\delta}, \boldsymbol{\phi}|y) \propto p(\boldsymbol{\phi}) \frac{1}{\vartheta} \prod_{j=1}^{J} \delta_j^{a_j-1} \prod_{d=0}^{1} \prod_{j=1}^{J} (\lambda_{dj})^{y_{dj}} \exp(-\lambda_{dj}) \tag{16}$$

Replacing the expression of λ_{dj} from (13), we have

$$\begin{aligned} p(\vartheta, \boldsymbol{\delta}, \boldsymbol{\phi}|y) \quad &\propto \quad \frac{p(\boldsymbol{\phi})}{\vartheta} \prod_{j=1}^{J} \left\{\vartheta \exp\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right)\right\}^{y_{1j}} \delta_j^{y_{+j}+a_j-1} \\ &\times \quad \exp\left[-\left(1 + \vartheta \exp\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right)\right) \delta_j\right] \end{aligned} \tag{17}$$

Integrating out δ_j from the above expression, we have

$$p(\vartheta, \boldsymbol{\phi}|y) \propto \frac{p(\boldsymbol{\phi})}{\vartheta} \prod_{j=1}^{J} \frac{\Gamma(y_{+j} + a_j)}{\left[1 + \vartheta \text{exp}\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right)\right]^{y_{+j}+a_j}} \left\{\vartheta \text{exp}\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right)\right\}^{y_{1j}}$$

Now, performing the transformation from ϑ to w yields expression (14).

(ii) First, we perform the transformation from $\boldsymbol{\delta}$ to $(\boldsymbol{\theta}, \psi)$, where $\psi = \sum_{j=1}^{J} \delta_j$. Thus, $\delta_j = \theta_j \psi$, $j = 1, ..., J$. The jacobian of transformation will be ψ^{J-1}.

Using this transformation in (16) and after some manipulation, we have

$$\begin{aligned} p(\vartheta, \boldsymbol{\theta}, \boldsymbol{\phi}, \psi|y) &\propto p(\boldsymbol{\phi})\psi^{y_{++}+a_{+}-1}\vartheta^{y_{1+}-1} \prod_{j=1}^{J} \theta_j^{y_{+j}+a_j-1} \text{exp}\left(\boldsymbol{\phi}' \sum_{j=1}^{J} y_{1j} \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right) \\ &\times \text{exp}\left[-\psi\left\{\sum_{j=1}^{J} \theta_j \left(1 + \vartheta \text{exp}\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right)\right)\right\}\right] \end{aligned} \tag{18}$$

Now, integrating (18) w.r.t ϑ we obtain

$$\begin{aligned} p(\boldsymbol{\theta}, \boldsymbol{\phi}, \psi|y) &\propto p(\boldsymbol{\phi}) \frac{\Gamma(y_{1+})}{\left[\psi \sum_{j=1}^{J} \theta_j \text{exp}\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt\right)\right]^{y_{1+}}} \psi^{y_{++}+a_{+}-1} \prod_{j=1}^{J} \theta_j^{y_{+j}+a_j-1} \\ &\times \text{exp}\left(\boldsymbol{\phi}' \sum_{j=1}^{J} y_{1j} \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt - \psi \sum_{j=1}^{J} \theta_j\right) \end{aligned} \tag{19}$$

Integration of (19) w.r.t ψ yields (15) after some minor manipulation.

(iii) The order in which $p(\vartheta, \boldsymbol{\delta}, \boldsymbol{\phi}|y)$ is integrated w.r.t the parameters does not make any difference in the marginal posterior density of $p(\boldsymbol{\phi})$. Thus, integration of $p(w, \boldsymbol{\phi}|y)$ w.r.t w or $p(\boldsymbol{\theta}, \boldsymbol{\phi}|y)$ w.r.t $\boldsymbol{\theta}$ will yield the same marginal posterior density $p(\boldsymbol{\phi}|y)$ of $\boldsymbol{\phi}$.

Remark:

1. As in Seaman and Richardson (2004), the assumption of existence and finiteness of $E\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_q(t)\Psi(t)dt\right)$ and $E\left(\boldsymbol{\phi}' \int_{-c_1}^{-c_2} Z_r(t)\Psi(t)dt\right)$ is automatically satisfied provided the prior density $p(\boldsymbol{\phi})$ ensures that $E(\boldsymbol{\phi})$ exists and is finite.

2. The posterior propriety of $p(\vartheta, \boldsymbol{\delta}, \boldsymbol{\phi}|y)$ in (17) can be shown in a similar way to that in Seaman and Richardson (2004).

3. As in Seaman and Richardson (2004), it can be shown in the context of a case control study that as a tends to 0, the log odds ratio parameters $\boldsymbol{\phi}$ from the retrospective model converge in distribution to the log odds ratio parameters from the limiting prospective model given by

$$Y_{1j} \sim \text{Bin}(Y_{+j}, p_j)(j = 1, ..., J); \qquad \log\left(\frac{p_j}{1-p_j}\right) = \omega + \int_{-c_1}^{-c_2} Z_j(t)\Psi(t)dt$$

where $\boldsymbol{\phi} \sim p(\boldsymbol{\phi})$, $p(\omega) \propto 1$ and $\boldsymbol{\theta} \sim \text{Dirichlet}(a, ..., a)$.

A.3: Calculations of M_i and Q_i matrices

Here we explain the details of the calculation of the M_i and Q_i matrices/vectors which appear in the disease risk model (5) and (7) in the main paper.
The exposure trajectory model is given by

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_1 a_{ij} + \sum_{l=1}^{k} \beta_{l+1}(a_{ij} - \tau_l)_+ + b_i a_{ij} + e_{ij} \\ &= \beta_0 + \beta_1(t_{ij} + a_i^d) + \sum_{l=1}^{k} \beta_{l+1}(t_{ij} + a_i^d - \tau_l)_+ + b_i(t_{ij} + a_i^d) + e_{ij} \\ &= \Phi(t_{ij} + a_i^d)'\boldsymbol{\beta} + b_i(t_{ij} + a_i^d) + e_{ij} \end{aligned} \tag{20}$$

where $\Phi(t_{ij} + a_i^d) = [1, (t_{ij} + a_i^d), (t_{ij} + a_i^d - \tau_1)_+, ..., (t_{ij} + a_i^d - \tau_K)_+]'$.

The prospective disease model is given by

$$\begin{aligned} P(D_i = 1 | X_i(t + a_i^d), -c_1 \leq t \leq -c_2) &= L\left(\alpha + \delta a_i^d + \int_{-c_1}^{-c_2} X_i(t + a_i^d)\gamma(t)dt\right) \\ &= L\left(\alpha + \delta a_i^d + \int_{-c_1}^{-c_2} \left(\Phi(t + a_i^d)'\boldsymbol{\beta} + b_i(t + a_i^d)\right) \Psi(t)'\boldsymbol{\phi} dt\right) \\ &= L\left(\alpha + \delta a_i^d + \boldsymbol{\beta}' M_i(c_1, c_2)\boldsymbol{\phi} + b_i Q_i(c_1, c_2)\boldsymbol{\phi}\right) \end{aligned} \tag{21}$$

where $M_i(c_1, c_2) = \int_{-c_1}^{-c_2} \Phi(t + a_i^d)\Psi(t)'dt$ and $Q_i(c_1, c_2) = \int_{-c_1}^{-c_2} (t + a_i^d)\Psi(t)'dt$.

Case I : One knot in the exposure trajectory, no knots in the influence function i.e $\Phi(t + a_i^d) = [1, (t + a_i^d), (t + a_i^d - \tau_1)_+]'$ and $\Psi(t) = (1, t)'$. Then,

$$\Phi(t + a_i^d)\Psi(t)' = \begin{pmatrix} 1 & t \\ t + a_i^d & t(t + a_i^d) \\ (t + a_i^d - \tau_1)_+ & t(t + a_i^d - \tau_1)_+ \end{pmatrix}.$$

and

$$(t+a_i^d)\Psi(t)' = \left((t+a_i^d) \quad t(t+a_i^d)\right).$$

$M_i(c_1,c_2)$ and $Q_i(c_1,c_2)$ can now be obtained by simply integrating out the elements of the above matrices over the interval $[-c_1,-c_2]$ i.e

$$\begin{aligned}\int_{-c_1}^{-c_2} dt &= c_1 - c_2, \quad \int_{-c_1}^{-c_2} t dt = \frac{c_2^2-c_1^2}{2}, \quad \int_{-c_1}^{-c_2}(t+a_i^d)dt = \frac{c_2^2-c_1^2}{2} + a_i^d(c_1-c_2),\\ \int_{-c_1}^{-c_2} t(t+a_i^d)dt &= \frac{c_1^3-c_2^3}{3} + a_i^d\frac{c_2^2-c_1^2}{2}.\end{aligned}$$

$$\begin{aligned}\int_{-c_1}^{-c_2}(t+a_i^d-\tau_1)_+ dt &= \int_{R_{i1}}^{T_{i1}}(t+a_i^d-\tau_1)dt = \frac{T_{i1}^2-R_{i1}^2}{2} + (a_i^d-\tau_1)(T_{i1}-R_{i1})\\ \int_{-c_1}^{-c_2} t(t+a_i^d-\tau_1)_+ dt &= \int_{R_{i1}}^{T_{i1}} t(t+a_i^d-\tau_1)dt = \frac{T_{i1}^3-R_{i1}^3}{3} + (a_i^d-\tau_1)\frac{T_{i1}^2-R_{i1}^2}{2}.\end{aligned}$$

where $R_{i1} = \max(-c_1, \tau_1 - a_i^d)$ and $T_{i1} = \max(-c_2, \tau_1 - a_i^d)$. T_{ij} takes care of the case of $-c_2$ being less than $(\tau_j - a_i^d)$. However, something like should be very rare.

Q_i will be the same as the 2nd row of M_i. Addition of more knots in the trajectory would just increase the number of rows of M_i. For example, if we add another knot, say, τ_2 in the trajectory model (such that $\tau_2 > \tau_1$), then,

$$\Phi(t+a_i^d)\Psi(t)' = \begin{pmatrix} 1 & t \\ t+a_i^d & t(t+a_i^d) \\ (t+a_i^d-\tau_1)_+ & t(t+a_i^d-\tau_1)_+ \\ (t+a_i^d-\tau_2)_+ & t(t+a_i^d-\tau_2)_+ \end{pmatrix}.$$

Q_i would however remain unchanged.

Case II : Let us add a knot, say κ_v, to the influence function $\gamma(t)$ i.e

$$\gamma(t) = \phi_1 + \phi_2 t + \phi_3(t-\kappa_v)_+$$

$\Phi(t+a_i^d)$ would remain the same as before (i.e 2 knots in the trajectory). Then

$$\Phi(t+a_i^d)\Psi(t)' = \begin{pmatrix} 1 & t & (t-\kappa_v)_+ \\ t+a_i^d & t(t+a_i^d) & (t+a_i^d)(t-\kappa_v)_+ \\ (t+a_i^d-\tau_1)_+ & t(t+a_i^d-\tau_1)_+ & (t+a_i^d-\tau_1)_+(t-\kappa_v)_+ \\ (t+a_i^d-\tau_2)_+ & t(t+a_i^d-\tau_2)_+ & (t+a_i^d-\tau_2)_+(t-\kappa_v)_+ \end{pmatrix}$$

and

$$(t+a_i^d)\Psi(t)' = \begin{pmatrix} t+a_i^d & t(t+a_i^d) & (t+a_i^d)(t-\kappa_v)_+ \end{pmatrix}$$

The calculations of M_i and Q_i would be similar as in Case I while

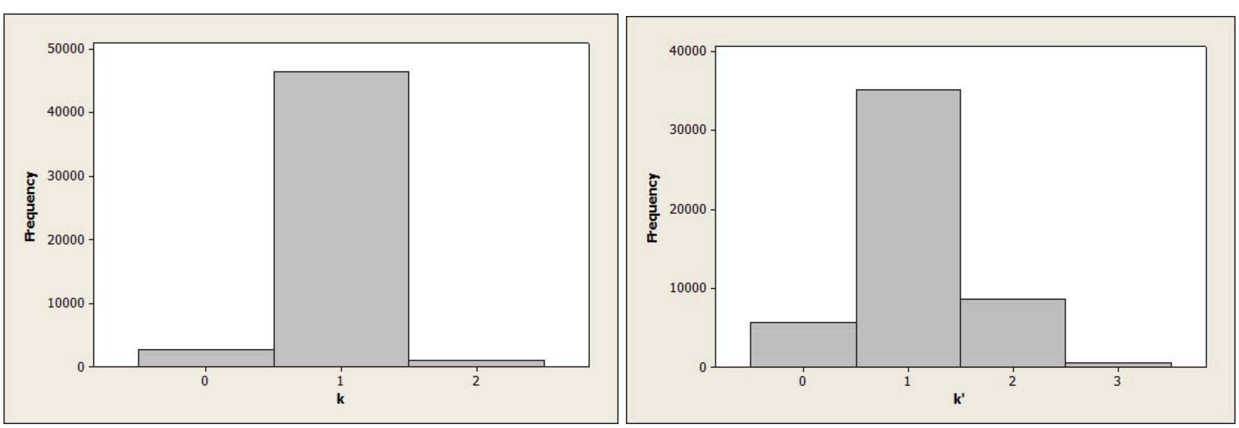
$$\begin{aligned}
\int_{-c_1}^{-c_2}(t-\kappa_v)_+dt &= \int_{S_{iv}}^{W_{iv}}(t-\kappa_v)dt = \frac{W_{iv}^2-S_{iv}^2}{2}-\kappa_v(W_{iv}-S_{iv}).\\
\int_{-c_1}^{-c_2}(t+a_i^d)(t-\kappa_v)_+dt &= \int_{S_{iv}}^{W_{iv}}(t+a_i^d)(t-\kappa_v)dt\\
&= \frac{W_{iv}^3-S_{iv}^3}{3}+(a_i^d-\kappa_v)\frac{W_{iv}^2-S_{iv}^2}{2}-a_i^d\kappa_v(W_{iv}-S_{iv}).\\
\int_{-c_1}^{-c_2}(t+a_i^d-\tau_u)_+(t-\kappa_v)_+dt &= \int_{S_{iuv}}^{W_{iuv}}(t+a_i^d-\tau_u)(t-\kappa_v)dt\\
&= \frac{W_{iuv}^3-S_{iuv}^3}{3}+(a_i^d-\tau_u-\kappa_v)\frac{W_{iuv}^2-S_{iuv}^2}{2}-(a_i^d-\tau_u)\kappa_v(W_{iuv}-S_{iuv})
\end{aligned}$$

where $S_{iv}=\max(-c_1,\kappa_v)$, $W_{iv}=\max(-c_2,\kappa_v)$, $S_{iuv}=\max\{-c_1,\max(\tau_u-a_i^d,\kappa_v)\}$ and $W_{iuv}=\max\{-c_2,\max(\tau_u-a_i^d,\kappa_v)\}$.

As before, Q_i would be the second row of M_i. Addition of more knots to the influence function would increase the columns of M_i and Q_i. The different elements of M_i can then be calculated as shown above.

A.4: Plots

In the following, Figure 1(a) and 1(b) depicts the posterior distribution of the number of knots for the exposure trajectory and the influence function while Figure 2 shows the smoothed density plots of the risk scores for the cases (R_1) and controls (R_0).



(a) Exposure Trajectory (b) Influence Function

Figure 1: Posterior distribution of the number of knots for the exposure trajectory and influence function.