Detection of age-specific genetic effects for age-at-onset complex traits

Introduction

The advent of genome-wide association (GWA) studies has enabled to better understand the genetic architecture and estimate its properties for many important complex traits. Among those are traits related to age-at-onset that stand out from others as they can almost always be right censored, meaning that we might only be able to measure the last known time without the event. Therefore, time-to-event traits have posed certain difficulties potentially leading to hampered statistical power or simply neglecting those traits from analyses.

Recently, the methodology for conducting GWA studies for time-to-event traits has greatly improved. The usage of martingale residuals has enabled combining censoring and timing information into one summary statistic which can be then plugged in for other available methods and software for further analysis [1, 2]. For joint marker effect estimation that combines censoring and timing informations in the partial likelihood in regularised regression framework there exists Cox-LASSO [3] implementation snpnet designed for variable selection and estimation in high-dimensional marker data [4,5]. To analyse time-to-event traits in mixed model framework the COXMEG method has recently been proposed [6], although it might not yet be completely scalable for biobank-scale data sets. Joint effect size estimation along with variable selection and effect size classification has been lately implemented within Bayesian framework in the BayesW model [7] that also enables partitioning genetic variance between different annotations. The latter approach has been shown to be efficient for genetic prediction and effect size classification enables better understanding of the genetic architecture.

Regardless of the advances proposed by the previous models, all of the previously described methods are assuming that the marker effect sizes on the trait remain constant throughout individual's life. That underlying assumption is definitely natural for many complex traits that are not describing age-at-onset. For example, once an individual has reached maximal growth, the height will remain relatively similar throughout the life and thus it is reasonable to model constant marker effect sizes having impacted height. In contrast to that, age-at-onset traits are affected by the hazard of onset that could vary throughout the life and it is possible that the changes in hazard are the result of markers having different effect throughout individual's life.

There exists substantial evidence that genetic effects on individual's phenotype can vary significantly throughout lifespan. For example, higher ages have been shown to have lower heritabilities across many continous complex traits [8].

Results

Discussion

Methods

Model outline

In the BayesW model lifespan T_i for an individual i follows a Weibull distribution such that $E(\log T_i|\mu, x_i, \beta, \alpha) = \mu + x_i\beta$ and $Var(\log T_i|\mu, x_i, \beta, \alpha) = \frac{\pi^2}{6\alpha^2}$. Therefore as a log-linear model it can be written down as

$$\log T_i = \mu + x_i\beta + \frac{w_i}{\alpha} + \frac{K}{\alpha}, \quad (1)$$

where K is the Euler-Mascheroni constant and w_i is an error term coming from the standard extreme value distribution. This model can be represented via hazard function $\lambda_i(t)$ for individual i by

$$\lambda_i(t) = \alpha t^{\alpha-1} \exp(-\alpha(\mu + x_i\beta) - K). \tag{2}$$

We will now expand the model by allowing different effect sizes throughout time. Suppose that we have two time intervals $[0, \tau)$ and $[\tau, \infty)$ that we will call epoch one and epoch two. Both of those epochs have their time-specific effect sizes that will be expressed only within those intervals. Lifespan T_i for individual i has to belong to either of those epochs. We will define the hazard function for the individual i in the following way

$$\lambda_i(t) = \begin{cases} \alpha t^{\alpha-1} \exp\bigl(-\alpha(\mu + x_i\beta^1) - K\bigr), & \text{if } t < \tau \\ \alpha t^{\alpha-1} \exp\bigl(-\alpha(\mu + x_i\beta^2) - K\bigr), & \text{if } t \geq \tau, \end{cases} \tag{3}$$

where β^1 and β^2 are vectors of M elements corresponding to M marker effects in first and second epoch respectively. Let's denote residual vectors for individual i in the following way $\varepsilon_i^1 = \log(t_i) - \mu - x_i\beta^1$, $\varepsilon_i^2 = \log(t_i) - \mu - x_i\beta^2$, $\varepsilon_i^3 = \log(\tau) - \mu - x_i\beta^1$ and $\varepsilon_i^4 = \log(\tau) - \mu - x_i\beta^2$. Given our definition of the hazard function we can find the corresponding survival function for individual i:

$$S_i(t) = \begin{cases} \exp\bigl[-\exp\bigl(\alpha\varepsilon_i^1 - K\bigr)\bigr], & \text{if } t < \tau \\ \exp\bigl[-\exp\bigl(\alpha\varepsilon_i^2 - K\bigr) - \exp\bigl(\alpha\varepsilon_i^3 - K\bigr) + \exp\bigl(\alpha\varepsilon_i^4 - K\bigr)\bigr], & \text{if } t \geq \tau. \end{cases} \tag{4}$$

The expression for the likelihood can be factorised into two parts

$$p(D|\Theta) = \prod_{i=1}^{n} (\lambda_i(t_i))^{d_i} S_i(t_i) = \prod_{i \in E_1} (\lambda_i(t_i))^{d_i} S_i(t_i) \prod_{i \in E_2} (\lambda_i(t_i))^{d_i} S_i(t_i) \tag{5}$$

where D denotes the data, Θ denotes all of the model parameters, d_i is the censoring indicator for individual i ($d_i = 1$ if event happened, 0 otherwise), E_q is a set of indices for individuals whose last known time without event belongs to epoch q.

The log-likelihood of the model is following:

$$\begin{aligned} \log(p(D|\Theta)) = d^1 \log\bigl(\alpha^1\bigr) + d^2 \log\bigl(\alpha^2\bigr) - Kd - d^1\alpha^1\mu - d^2\alpha^2\mu + \\ (\alpha^1 - 1) \sum_{i \in E_1} \log(t_i) d_i - \alpha^1 \sum_{i \in E_1} d_i x_i \beta^1 - \sum_{i \in E_1} \exp\bigl(\alpha^1 \varepsilon_i^1 - K\bigr) + \\ (\alpha^2 - 1) \sum_{i \in E_2} \log(t_i) d_i - \alpha^2 \sum_{i \in E_2} d_i x_i \beta^2 - \sum_{i \in E_2} \exp\bigl(\alpha^2 \varepsilon_i^2 - K\bigr) + \\ \sum_{i \in E_2} \exp\bigl(\alpha^2 \varepsilon_i^4 - K\bigr) - \sum_{i \in E_2} \exp\bigl(\alpha^1 \varepsilon_i^3 - K\bigr), \end{aligned} \tag{6}$$

where d^1, d^2 and d are the numbers of uncensored individuals in epoch 1, epoch 2 and in total.

We will fix the following priors for the parameters. Let the prior distribution of α^q be a gamma distribution with parameters α_0, κ_0

$$p(\alpha^q) \propto (\alpha^q)^{\alpha_0 - 1} \exp(-\kappa_0 \alpha^q), \tag{7}$$

the prior for μ be normal with variance parameter σ_μ^2:

$$p(\mu) \propto \exp\biggl(-\frac{1}{2\sigma_\mu^2}\mu^2\biggr). \tag{8}$$

The rest of the parameters are epoch-specific. Suppose that the effects for marker j for epochs one and two follow a bivariate normal distribution

$$(\beta_j^1, \beta_j^2)|\sigma_{G1}^2, \sigma_{G2}^2, \gamma_j^1 = k, \gamma_j^2 = l \sim N\left(0, \Sigma_{kl}\right), j = 1, \dots, M \tag{9}$$

where

$$\Sigma_{kl} = \begin{bmatrix} \sigma_{G1}^2 C_{k1} & \rho\sigma_{G1}\sigma_{G2}\sqrt{C_{k1}C_{l2}} \\ \rho\sigma_{G1}\sigma_{G2}\sqrt{C_{k1}C_{l2}} & \sigma_{G2}^2 C_{l2} \end{bmatrix} \tag{10}$$

and here we assume that C_{k1} and C_{l2} are positive. Thus, given that the mixture components C_{k1} and C_{l2} are positive and finding the conditional distributions based on equation 9, the prior distributions for the effect sizes are

$$\beta_j^1|\beta_j^2, \sigma_{G1}^2, \sigma_{G2}^2, \gamma_j^1 = k, \gamma_j^2 = l \sim N\left(\rho\frac{\sigma_{G1}\sqrt{C_{k1}}}{\sigma_{G2}\sqrt{C_{l2}}}\beta_j^2, \sigma_{G1}^2 C_{k1}(1-\rho^2)\right) \tag{11}$$

$$\beta_j^2|\beta_j^1, \sigma_{G1}^2, \sigma_{G2}^2, \gamma_j^1 = k, \gamma_j^2 = l \sim N\left(\rho\frac{\sigma_{G2}\sqrt{C_{l2}}}{\sigma_{G1}\sqrt{C_{k1}}}\beta_j^1, \sigma_{G2}^2 C_{l2}(1-\rho^2)\right). \tag{12}$$

However, it can happen that when we evaluate effect for one epoch, then the effect size in the other epoch has been set to 0. For this situation we will define the corresponding priors using the following normal distribution if $\beta_j^2 = 0$ and $\gamma_j^2 = 0$:

$$\beta_j^1|\beta_j^2 = 0, \sigma_{G1}^2, \gamma_j^1 = k, \gamma_j^2 = 0 \sim N\left(0, \sigma_{G1}^2 C_{k1}(1-\rho^2)\right) \tag{13}$$

and analogously if $\beta_j^1 = 0$ (and thus $\gamma_j^1 = 0$) then we define our prior for β_j^1 as

$$\beta_j^2|\beta_j^1 = 0, \sigma_{G2}^2, \gamma_j^2 = k, \gamma_j^1 = 0 \sim N\left(0, \sigma_{G2}^2 C_{l2}(1-\rho^2)\right). \tag{14}$$

The prior for $\gamma_j^q, q \in \{1, 2\}$ is multinomial:

$$p(\gamma_j^q|\pi^q) = \pi_{0q}^{I(\gamma_j^q=0)} \cdot \ldots \cdot \pi_{Lq}^{I(\gamma_j^q=L)}, \tag{15}$$

the prior probabilities of belonging to each of the mixture distributions k are stored in $L+1$-dimensional vector π^q with the prior for π^q a Dirichlet distribution

$$p(\pi^q) = \text{Dirichlet}(\mathbf{p}_\text{L}), \tag{16}$$

where $I(\cdot)$ is the indicator function and $\mathbf{p}_\text{L}$ is the $L+1$-dimensional vector with prior values.

$$\begin{aligned}\log p(\mu|D, \alpha, \beta^1, \beta^2) = \text{const} - d\alpha\mu - \sum_{i\in E_1}\exp\bigl(\alpha\varepsilon_i^1 - K\bigr) - \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^2 - K\bigr) + \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^4 - K\bigr) \\ - \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^3 - K\bigr) - \frac{1}{2\sigma_\mu^2}\mu^2\end{aligned} \tag{17}$$

$$\begin{aligned}\log p(\mu|D, \alpha^1, \alpha^2, \beta^1, \beta^2) = \text{const} - d^1\alpha^1\mu - d^2\alpha^2\mu - \sum_{i\in E_1}\exp\bigl(\alpha^1\varepsilon_i^1 - K\bigr) - \sum_{i\in E_2}\exp\bigl(\alpha^2\varepsilon_i^2 - K\bigr) + \sum_{i\in E_2}\exp\bigl(\alpha^2\varepsilon_i^4 - K\bigr) \\ - \sum_{i\in E_2}\exp\bigl(\alpha^1\varepsilon_i^3 - K\bigr) - \frac{1}{2\sigma_\mu^2}\mu^2\end{aligned} \tag{18}$$

$$\log p(\mu^1|D,\alpha^1,\beta^1) \;=\; \text{const} \;-\; d^1\alpha^1\mu^1 \;-\; \sum_{i\in E_1}\exp\big(\alpha^1\varepsilon_i^1 - K\big) \;-\; \sum_{i\in E_2}\exp\big(\alpha^1\varepsilon_i^3 - K\big) \;-\; \frac{1}{2\sigma_\mu^2}\mu^2 \quad (19)$$

$$\log p(\mu^2|D,\alpha^2,\beta^2) \;=\; \text{const} \;-\; d^2\alpha^2\mu^2 \;-\; \sum_{i\in E_2}\exp\big(\alpha^2\varepsilon_i^2 - K\big) \;+\; \sum_{i\in E_2}\exp\big(\alpha^2\varepsilon_i^4 - K\big) \;-\; \frac{1}{2\sigma_\mu^2}\mu^2 \quad (20)$$

$$\begin{aligned}\log p(\alpha|D,\mu,\beta^1,\beta^2) &= \text{const}+(d+\alpha_0-1)\log(\alpha)-d\alpha\mu-\kappa_0\alpha+\alpha\sum_{i\in E_1}\log(t_i)d_i-\alpha\sum_{i\in E_1}d_ix_i\beta^1-\sum_{i\in E_1}\exp\big(\alpha\varepsilon_i^1-K\big)\\ &+\alpha\sum_{i\in E_2}\log(t_i)d_i-\alpha\sum_{i\in E_2}d_ix_i\beta^2-\sum_{i\in E_2}\exp\big(\alpha\varepsilon_i^2-K\big)+\sum_{i\in E_2}\exp\big(\alpha\varepsilon_i^4-K\big)-\sum_{i\in E_2}\exp\big(\alpha\varepsilon_i^3-K\big)\\ &= \text{const}+(d+\alpha_0-1)\log(\alpha)+\alpha\sum_{i\in E_1}d_i[\log(t_i)-\mu-x_i\beta^1]+\alpha\sum_{i\in E_2}d_i[\log(t_i)-\mu-x_i\beta^2]+\\ &\exp(-K)\left[-\sum_{i\in E_1}\exp\big(\alpha\varepsilon_i^1\big)-\sum_{i\in E_2}\exp\big(\alpha\varepsilon_i^2\big)+\sum_{i\in E_2}\exp\big(\alpha\varepsilon_i^4\big)-\sum_{i\in E_2}\exp\big(\alpha\varepsilon_i^3\big)\right]-\kappa_0\alpha \quad (21)\end{aligned}$$

$$\begin{aligned}\log p(\alpha^1|D,\mu,\beta^1,\beta^2) &= \text{const}+(d^1+\alpha_0-1)\log\big(\alpha^1\big)-d^1\alpha^1\mu-\kappa_0\alpha^1+\alpha^1\sum_{i\in E_1}\log(t_i)d_i-\alpha^1\sum_{i\in E_1}d_ix_i\beta^1\\ &-\sum_{i\in E_1}\exp\big(\alpha^1\varepsilon_i^1-K\big)+\sum_{i\in E_2}\exp\big(\alpha^1\varepsilon_i^3-K\big)\\ &= \text{const}+(d^1+\alpha_0-1)\log\big(\alpha^1\big)+\alpha^1\sum_{i\in E_1}d_i[\log(t_i)-\mu-x_i\beta^1]+\\ &\exp(-K)\left[-\sum_{i\in E_1}\exp\big(\alpha^1\varepsilon_i^1\big)-\sum_{i\in E_2}\exp\big(\alpha^1\varepsilon_i^3\big)\right]-\kappa_0\alpha^1 \quad (22)\end{aligned}$$

$$\begin{aligned}\log p(\alpha^2|D,\mu,\beta^1,\beta^2) &= \text{const}+(d^2+\alpha_0-1)\log\big(\alpha^2\big)-d^2\alpha^2\mu-\kappa_0\alpha^2\\ &+\alpha^2\sum_{i\in E_2}\log(t_i)d_i-\alpha^2\sum_{i\in E_2}d_ix_i\beta^2-\sum_{i\in E_2}\exp\big(\alpha^2\varepsilon_i^2-K\big)+\sum_{i\in E_2}\exp\big(\alpha^2\varepsilon_i^4-K\big)\\ &= \text{const}+(d^2+\alpha_0-1)\log\big(\alpha^2\big)+\alpha^2\sum_{i\in E_2}d_i[\log(t_i)-\mu-x_i\beta^2]+\\ &\exp(-K)\left[-\sum_{i\in E_2}\exp\big(\alpha^2\varepsilon_i^2\big)+\sum_{i\in E_2}\exp\big(\alpha^2\varepsilon_i^4\big)\right]-\kappa_0\alpha^2 \quad (23)\end{aligned}$$

$$\begin{aligned}\log p(\beta_j^1|D,\mu,\alpha,\sigma_{G1}^2,\sigma_{G2}^2,\beta_j^2,\gamma_j^1=k,\gamma_j^2=l\neq 0) &= \text{const}-\alpha\beta_j^1\sum_{i\in E_1}d_ix_{ij}\\ -\sum_{i\in E_1}\exp\big(\alpha\varepsilon_i^1-K\big)-\sum_{i\in E_2}\exp\big(\alpha\varepsilon_i^3-K\big)&-\frac{1}{2\sigma_{G1}^2C_{k1}(1-\rho^2)}\left(\beta_j^1-\rho\frac{\sigma_{G1}\sqrt{C_{k1}}}{\sigma_{G2}\sqrt{C_{l2}}}\beta_j^2\right)^2. \quad (24)\end{aligned}$$

$$\log p(\beta_j^1|D,\mu,\alpha,\sigma_{G1}^2,\sigma_{G2}^2,\beta_j^2,\gamma_j^1=k,\gamma_j^2=0) = \text{const} - \alpha\beta_j^1\sum_{i\in E_1} d_i x_{ij} \\ - \sum_{i\in E_1}\exp\bigl(\alpha\varepsilon_i^1 - K\bigr) - \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^3 - K\bigr) - \frac{\bigl(\beta_j^1\bigr)^2}{2\sigma_{G1}^2 C_{k1}(1-\rho^2)}. \quad (25)$$

$$\log p(\beta_j^2|D,\mu,\alpha,\sigma_{G1}^2,\sigma_{G2}^2,\beta_j^1,\gamma_j^1=k,\gamma_j^2=l\neq 0) = \text{const} - \alpha\beta_j^2\sum_{i\in E_2} d_i x_{ij} \\ - \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^2 - K\bigr) + \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^4 - K\bigr) - \frac{1}{2\sigma_{G2}^2 C_{l2}(1-\rho^2)}\left(\beta_j^2 - \rho\frac{\sigma_{G2}\sqrt{C_{l2}}}{\sigma_{G1}\sqrt{C_{k1}}}\beta_j^1\right)^2. \quad (26)$$

$$\log p(\beta_j^2|D,\mu,\alpha,\sigma_{G1}^2,\sigma_{G2}^2,\beta_j^1,\gamma_j^1=k,\gamma_j^2=0) = \text{const} - \alpha\beta_j^2\sum_{i\in E_2} d_i x_{ij} \\ - \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^2 - K\bigr) + \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^4 - K\bigr) - \frac{\bigl(\beta_j^2\bigr)^2}{2\sigma_{G2}^2 C_{l2}(1-\rho^2)}. \quad (27)$$

$$\log p(\delta_j|D,\mu,\alpha,\beta^1,\beta^2,\delta^1,\delta^2) = \text{const} - \alpha\delta_j\left(\sum_{i\in E_1} z_{ij}d_i + \sum_{i\in E_2} z_{ij}d_i\right) - \sum_{i\in E_1}\exp\bigl(\alpha\varepsilon_i^1 - K\bigr) \\ - \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^2 - K\bigr) + \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^4 - K\bigr) - \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^3 - K\bigr) - \frac{1}{2\sigma_\delta^2}\delta^2 \quad (28)$$

$$\log p(\delta_j^1|D,\mu,\alpha,\beta^1,\beta^2,\delta,\delta^2) = \text{const} - \alpha\delta_j^1\sum_{i\in E_1} z_{ij}^1 d_i - \sum_{i\in E_1}\exp\bigl(\alpha\varepsilon_i^1 - K\bigr) - \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^3 - K\bigr) - \frac{1}{2\sigma_\delta^2}\delta^2 \quad (29)$$

$$\log p(\delta_j^2|D,\mu,\alpha,\beta^1,\beta^2,\delta^1,\delta) = \text{const} - \alpha\delta_j^2\sum_{i\in E_2} z_{ij}^2 d_i - \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^2 - K\bigr) + \sum_{i\in E_2}\exp\bigl(\alpha\varepsilon_i^4 - K\bigr) - \frac{1}{2\sigma_\delta^2}\delta^2 \quad (30)$$

We define $\eta_{kl}^1 := \begin{cases}\rho\frac{\sigma_{G1}\sqrt{C_{k1}}}{\sigma_{G2}\sqrt{C_{l2}}}\beta_j^2, & l\neq 0\\ 0, & l=0\end{cases}$ and $\eta_{kl}^2 := \begin{cases}\rho\frac{\sigma_{G2}\sqrt{C_{l2}}}{\sigma_{G1}\sqrt{C_{k1}}}\beta_j^1, & k\neq 0\\ 0, & k=0\end{cases}$.

$$\hat\sigma_{k1} = \frac{1}{\sqrt{2}\sqrt{1+\alpha^2 C_{k1}(1-\rho^2)\sigma_{G1}^2\left(\sum_{i\in E_1} x_{ij}^2\exp(u_i^1 - \alpha x_{ij}\eta_{kl}^1) + \sum_{i\in E_2} x_{ij}^2\exp(u_i^3 - \alpha x_{ij}\eta_{kl}^1)\right)}}$$

$$\hat\sigma_{l2} = \frac{1}{\sqrt{2}\sqrt{1+\alpha^2 C_{l2}(1-\rho^2)\sigma_{G2}^2\left(\sum_{i\in E_2} x_{ij}^2\exp(u_i^2 - \alpha x_{ij}\eta_{kl}^2) - \sum_{i\in E_2} x_{ij}^2\exp(u_i^4 - \alpha x_{ij}\eta_{kl}^2)\right)}}$$

$$\frac{g_k^1(\hat{\sigma}_{k1}\sqrt{2}t_r)}{T_1} = \exp\Bigg\{ -\alpha \sum_{i\in E_1} d_i x_{ij}(\hat{\sigma}_{k1}\sqrt{2}t_r\sqrt{2\sigma_{G1}^2 C_{k1}(1-\rho^2)} + \eta_{kl}^1) + \sum_{i\in E_1}\Big[-\exp\Big(u_i^1 - \alpha x_{ij}(\hat{\sigma}_{k1}\sqrt{2}t_r\sqrt{2\sigma_{G1}^2 C_{k1}(1-\rho^2)} + \eta_{kl}^1)\Big) + \exp\big(u_i^1\big)\Big] + \sum_{i\in E_2}\Big[-\exp\Big(u_i^3 - \alpha x_{ij}(\hat{\sigma}_{k1}\sqrt{2}t_r\sqrt{2\sigma_{G1}^2 C_{k1}(1-\rho^2)} + \eta_{kl}^1)\Big) + \exp\big(u_i^3\big)\Big] - (\hat{\sigma}_{k1}\sqrt{2}t_r)^2 \Bigg\}$$

$$\frac{g_k^2(\hat{\sigma}_{l2}\sqrt{2}t_r)}{T_2} = \exp\Bigg\{ -\alpha \sum_{i\in E_2} d_i x_{ij}(\hat{\sigma}_{l2}\sqrt{2}t_r\sqrt{2\sigma_{G2}^2 C_{l2}(1-\rho^2)} + \eta_{kl}^2) + \sum_{i\in E_2}\Big[-\exp\Big(u_i^2 - \alpha x_{ij}(\hat{\sigma}_{k1}\sqrt{2}t_r\sqrt{2\sigma_{G2}^2 C_{l2}(1-\rho^2)} + \eta_{kl}^2)\Big) + \exp\big(u_i^2\big)\Big] + \sum_{i\in E_2}\Big[\exp\Big(u_i^4 - \alpha x_{ij}(\hat{\sigma}_{l2}\sqrt{2}t_r\sqrt{2\sigma_{G2}^2 C_{l2}(1-\rho^2)} + \eta_{kl}^2)\Big) - \exp\big(u_i^4\big)\Big] - (\hat{\sigma}_{l2}\sqrt{2}t_r)^2 \Bigg\}$$

Derivation for the sparse calculations 72

We define $M_j = \exp\Big(-\alpha\eta_{kl}^1 \frac{1}{s_j}\Big)$ and $N_j = M_j^{-\bar{\xi}_j} = \exp\Big(\alpha\eta_{kl}^1 \frac{\bar{\xi}_j}{s_j}\Big)$.

$$\hat{\sigma}_{k1} = \frac{1}{\sqrt{2}}\Bigg[1 + \alpha^2 C_{k1}(1-\rho^2)\sigma_{G1}^2 \frac{1}{s_j^2} N_j\Bigg((\bar{\xi}_j)^2\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=0}}\exp\big(u_i^1\big) + \sum_{\substack{i\in E_2\\ \xi_{ij}=0}}\exp\big(u_i^3\big)\Bigg) + M_j\Bigg((1-\bar{\xi}_j)^2\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=1}}\exp\big(u_i^1\big) + \sum_{\substack{i\in E_2\\ \xi_{ij}=1}}\exp\big(u_i^3\big)\Bigg) + (2-\bar{\xi}_j)^2 M_j\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=2}}\exp\big(u_i^1\big) + \sum_{\substack{i\in E_2\\ \xi_{ij}=2}}\exp\big(u_i^3\big)\Bigg)\Bigg)\Bigg)\Bigg]^{-0.5}$$

Let us define $B_j = \exp\Big(\alpha\frac{\bar{\xi}_j}{s_j}(s\sqrt{2C_{k1}\sigma_{G1}^2(1-\rho^2)} + \eta_{kl}^1)\Big)$ and $C_j = \exp\Big(-\frac{\alpha}{s_j}(s\sqrt{2C_{k1}\sigma_{G1}^2(1-\rho^2)} + \eta_{kl}^1)\Big)$.

$$\sum_{i\in E_1}\exp\Big(u_i^1 - \alpha x_{ij}(\hat{\sigma}_{k1}\sqrt{2}t_r\sqrt{2\sigma_{G1}^2 C_{k1}(1-\rho^2)} + \eta_{kl}^1)\Big) + \sum_{i\in E_2}\exp\Big(u_i^3 - \alpha x_{ij}(\hat{\sigma}_{k1}\sqrt{2}t_r\sqrt{2\sigma_{G1}^2 C_{k1}(1-\rho^2)} + \eta_{kl}^1)\Big) = B_j\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=0}}\exp\big(u_i^1\big) + \sum_{\substack{i\in E_2\\ \xi_{ij}=0}}\exp\big(u_i^3\big) + C_j\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=1}}\exp\big(u_i^1\big) + \sum_{\substack{i\in E_2\\ \xi_{ij}=1}}\exp\big(u_i^3\big) + C_j\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=2}}\exp\big(u_i^1\big) + \sum_{\substack{i\in E_2\\ \xi_{ij}=2}}\exp\big(u_i^3\big)\Bigg)\Bigg)\Bigg)$$

To rewrite equations 24 and 27, we define $H_j = \exp\left(\frac{-\alpha\beta_j^1}{s_j}\right)$ and $G_j = H_j^{-\bar{\xi}_j} = \exp\left(\frac{\alpha\bar{\xi}_j\beta_j^1}{s_j}\right)$.

$$\sum_{i\in E_1}\exp\big(\alpha(\log(t_i) - \mu - x_i\beta^1) - K\big) + \sum_{i\in E_2}\exp\big(\alpha(\log(\tau) - \mu - x_i\beta^1) - K\big) = = G_j\Bigg[\sum_{\substack{i\in E_1\\ \xi_{ij}=0}}\exp\big(u_i^1\big) + \sum_{\substack{i\in E_2\\ \xi_{ij}=0}}\exp\big(u_i^3\big) + H_j\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=1}}\exp\big(u_i^1\big) + \sum_{\substack{i\in E_2\\ \xi_{ij}=1}}\exp\big(u_i^3\big) + H_j\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=2}}\exp\big(u_i^1\big) + \sum_{\substack{i\in E_2\\ \xi_{ij}=2}}\exp\big(u_i^3\big)\Bigg)\Bigg)\Bigg].$$

Three epoch model outline

Let's denote residual vectors for individual i in the following way

$$\begin{aligned}
\varepsilon_i^{(+j)} &= \log(\tau_j) - \mu - x_i\beta^{(j)}, \quad j = 1, 2.\\
\varepsilon_i^{(-j)} &= \log(\tau_j) - \mu - x_i\beta^{(j+1)}, \quad j = 1, 2.\\
\varepsilon_i^{(j)} &= \log(t_i) - \mu - x_i\beta^{(j)}, \quad j = 1, 2, 3.
\end{aligned}$$

Let's define an artificial epoch point $\tau_0 := 0$. We can find the corresponding survival function for individual i:

$$S_i(t) = \sum_{k=1}^{3} \exp\left[-\sum_{j=1}^{k-1} \exp\Big(\alpha\varepsilon_i^{(+j)} - K\Big) - \exp\Big(\alpha\varepsilon_i^{(k)} - K\Big) + \sum_{j=1}^{k-1} \exp\Big(\alpha\varepsilon_i^{(-j)} - K\Big)\right] \cdot \mathbb{1}_{\tau_{k-1} \le t < \tau_k} \qquad (31)$$

The log-likelihood of the model is the following:

$$\begin{aligned}
\log(p(D|\Theta)) = d\log(\alpha) - \alpha\mu d - Kd + (\alpha - 1)\sum_{k=1}^{3}\sum_{i\in E_k} d_i \log(t_i) - \alpha \sum_{k=1}^{3}\sum_{i\in E_k} d_i x_i \beta^{(k)}\\
+ \exp(-K)\sum_{k=1}^{3}\sum_{i\in E_k}\left[-\sum_{j=1}^{k-1}\exp\Big(\alpha\varepsilon_i^{(+j)}\Big) - \exp\Big(\alpha\varepsilon_i^{(k)}\Big) + \sum_{j=1}^{k-1}\exp\Big(\alpha\varepsilon_i^{(-j)}\Big)\right],
\end{aligned}$$

where d is the number of uncensored individuals. We assume the same priors as in (7), (8), (15) and (16) while for $\beta^{(k)}$'s we introduce the following conditional prior:

$$(\beta_j^{(1)}, \beta_j^{(2)}, \beta_j^{(3)})|(\sigma_{Gk}^2)_{k=1}^3, \gamma_j^1 = k, \gamma_j^{(2)} = l, \gamma_j^{(2)} = m \sim N\left(\mathbf{0}, \Sigma_{klm}\right), j = 1, \dots, M \qquad (32)$$

where

$$\Sigma_{klm} = \begin{bmatrix} \sigma_{G1}^2 C_{k1} & \rho_{12}\sigma_{G1}\sigma_{G2}\sqrt{C_{k1}C_{l2}} & \rho_{13}\sigma_{G1}\sigma_{G3}\sqrt{C_{k1}C_{m3}}\\ \rho_{12}\sigma_{G1}\sigma_{G2}\sqrt{C_{k1}C_{l2}} & \sigma_{G2}^2 C_{l2} & \rho_{23}\sigma_{G2}\sigma_{G3}\sqrt{C_{l2}C_{m3}}\\ \rho_{13}\sigma_{G1}\sigma_{G3}\sqrt{C_{k1}C_{m3}} & \rho_{23}\sigma_{G2}\sigma_{G3}\sqrt{C_{l2}C_{m3}} & \sigma_{G3}^2 C_{m3}\end{bmatrix} \qquad (33)$$

Note that here we assume that C_{k1}, C_{l2} and C_{m3} are positive. Thus, given that the mixture components C_{k1}, C_{l2} and C_{m3} are positive we find the conditional distributions based on equation (32). Therefore, the prior distributions for the effect sizes are

$$\begin{aligned}
&\beta_j^{(1)}|\beta_j^{(2)}, \beta_j^{(3)}(\sigma_{Gk}^2)_{k=1}^3(\gamma_j^{(k)})_{k=1}^3 = (k, l, m)\\
&\sim N\left(\frac{\sigma_{G1}\sqrt{C_{k1}}(\rho_{13} - \rho_{12}\rho_{23})\beta_j^{(3)}}{\sigma_{G3}\sqrt{C_{m3}}(1 - \rho_{23}^2)} + \frac{\sigma_{G1}\sqrt{C_{k1}}(\rho_{12} - \rho_{13}\rho_{23})\beta_j^{(2)}}{\sigma_{G2}\sqrt{C_{l2}}(1 - \rho_{23}^2)}, \sigma_{G1}^2 C_{k1}\left[1 - \frac{\rho_{12}^2 + \rho_{13}^2 - 2\rho_{12}\rho_{23}\rho_{13}}{1 - \rho_{23}^2}\right]\right)
\end{aligned}$$

$$\begin{aligned}
&\beta_j^{(2)}|\beta_j^{(1)}, \beta_j^{(3)}(\sigma_{Gk}^2)_{k=1}^3(\gamma_j^{(k)})_{k=1}^3 = (k, l, m)\\
&\sim N\left(\frac{\sigma_{G2}\sqrt{C_{l2}}(\rho_{23} - \rho_{12}\rho_{13})\beta_j^{(3)}}{\sigma_{G3}\sqrt{C_{m3}}(1 - \rho_{13}^2)} + \frac{\sigma_{G2}\sqrt{C_{l2}}(\rho_{12} - \rho_{13}\rho_{23})\beta_j^{(1)}}{\sigma_{G1}\sqrt{C_{k1}}(1 - \rho_{13}^2)}, \sigma_{G2}^2 C_{l2}\left[1 - \frac{\rho_{12}^2 + \rho_{23}^2 - 2\rho_{12}\rho_{13}\rho_{23}}{1 - \rho_{13}^2}\right]\right)
\end{aligned}$$

$$\begin{aligned}
&\beta_j^{(3)}|\beta_j^{(1)}, \beta_j^{(2)}(\sigma_{Gk}^2)_{k=1}^3(\gamma_j^{(k)})_{k=1}^3 = (k, l, m)\\
&\sim N\left(\frac{\sigma_{G3}\sqrt{C_{m3}}(\rho_{23} - \rho_{13}\rho_{12})\beta_j^{(2)}}{\sigma_{G2}\sqrt{C_{l2}}(1 - \rho_{12}^2)} + \frac{\sigma_{G3}\sqrt{C_{m3}}(\rho_{13} - \rho_{12}\rho_{23})\beta_j^{(1)}}{\sigma_{G1}\sqrt{C_{k1}}(1 - \rho_{12}^2)}, \sigma_{G3}^2 C_{m3}\left[1 - \frac{\rho_{13}^2 + \rho_{23}^2 - 2\rho_{13}\rho_{12}\rho_{23}}{1 - \rho_{12}^2}\right]\right)
\end{aligned}$$

However, it can happen that when we evaluate effect for one epoch, then the effect size in at least one of the other epochs has been set to 0. For this situation we will define the corresponding priors by using the equations written above, but with the following agreement: if $\beta_j^{(k)} = 0$ and $\gamma_j^{(k)} = 0$, then just ignore all the terms which contain $\beta_j^{(k)}$.

Note that the correlation coefficient in the bivariate normal distribution $(\beta_j^{(1)}, \beta_j^{(3)})|\beta_j^{(2)}$ is given by the partial correlation $\rho_{13.2} = \frac{\rho_{13}-\rho_{12}\rho_{23}}{\sqrt{(1-\rho_{12}^2)(1-\rho_{23}^2)}}$. So, to include in our model the fact that $\beta_j^{(1)}$ and $\beta_j^{(3)}$ should be conditionally independent given $\beta_j^{(2)}$ we impose a constraint $\rho_{13} := \rho_{12}\rho_{23}$. Hence, the previously mentioned priors became

$$\beta_j^{(1)}|\beta_j^{(2)}, \beta_j^{(3)}, (\sigma_{Gk}^2)_{k=1}^3, (\gamma_j^{(k)})_{k=1}^3 = (k,l,m) \sim N\left(\frac{\sigma_{G1}\sqrt{C_{k1}}\rho_{12}\beta_j^{(2)}}{\sigma_{G2}\sqrt{C_{l2}}}, \sigma_{G1}^2 C_{k1}\left[1 - \frac{\rho_{12}^2 - \rho_{12}^2\rho_{23}^2}{1-\rho_{23}^2}\right] = \sigma_{G1}^2 C_{k1}(1-\rho_{12}^2)\right)$$

$$\begin{aligned}&\beta_j^{(2)}|\beta_j^{(1)}, \beta_j^{(3)}(\sigma_{Gk}^2)_{k=1}^3(\gamma_j^{(k)})_{k=1}^3 = (k,l,m) \\ &\quad\sim N\left(\frac{\sigma_{G2}\sqrt{C_{l2}}\rho_{23}(1-\rho_{12}^2)\beta_j^{(3)}}{\sigma_{G3}\sqrt{C_{m3}}(1-\rho_{12}^2\rho_{23}^2)} + \frac{\sigma_{G2}\sqrt{C_{l2}}\rho_{12}(1-\rho_{23}^2)\beta_j^{(1)}}{\sigma_{G1}\sqrt{C_{k1}}(1-\rho_{12}^2\rho_{23}^2)}, \sigma_{G2}^2 C_{l2}\left[1 - \frac{\rho_{12}^2 + \rho_{23}^2 - 2\rho_{12}^2\rho_{23}^2}{1-\rho_{12}^2\rho_{23}^2}\right]\right)\end{aligned}$$

$$\beta_j^{(3)}|\beta_j^{(1)}, \beta_j^{(2)}(\sigma_{Gk}^2)_{k=1}^3(\gamma_j^{(k)})_{k=1}^3 = (k,l,m) \sim N\left(\frac{\sigma_{G3}\sqrt{C_{m3}}\rho_{23}\beta_j^{(2)}}{\sigma_{G2}\sqrt{C_{l2}}}, \sigma_{G3}^2 C_{m3}\left[1 - \frac{\rho_{23}^2 - \rho_{12}^2\rho_{23}^2}{1-\rho_{12}^2}\right] = \sigma_{G3}^2 C_{m3}(1-\rho_{23}^2)\right)$$

We continue with the list of posteriors:

$$\begin{aligned}\log p(\mu|D, \alpha, (\beta^{(k)})_{k=1}^3) = \text{const} + \exp(-K)\sum_{k=1}^3\sum_{i\in E_k}\left[-\sum_{j=1}^{k-1}\exp\Big(\alpha\varepsilon_i^{(+j)}\Big) - \exp\Big(\alpha\varepsilon_i^{(k)}\Big) + \sum_{j=1}^{k-1}\exp\Big(\alpha\varepsilon_i^{(-j)}\Big)\right] \\ - \alpha\mu d - \frac{1}{2\sigma_\mu^2}.\end{aligned}$$

$$\begin{aligned}\log p(\alpha|\mu, (\beta^{(k)})_{k=1}^3) = \text{const} + \alpha\sum_{k=1}^3\sum_{i\in E_k} d_i[\log(t_i) - \mu - x_i\beta^{(k)}] + (\alpha_0 + d - 1)\log(\alpha) - \kappa_0\alpha \\ + \exp(-K)\Big(-\sum_{j=1}^{k-1}\exp\Big(\alpha\varepsilon_i^{(+j)}\Big) - \exp\Big(\alpha\varepsilon_i^{(k)}\Big) + \sum_{j=1}^{k-1}\exp\Big(\alpha\varepsilon_i^{(-j)}\Big)\Big)\end{aligned}$$

$$\begin{aligned}\log p(\beta_j^{(1)}|D, \mu, \alpha, (\beta^{(k)})_{k=2,3}, (\sigma_{Gk}^2)_{k=1}^3, (\gamma_j^{(k)})_{k=1}^3 = (k,l,m)) = -\alpha\beta_j^{(1)}\sum_{i\in E_1} d_i x_{ij} \\ + \exp(-K)\Big(-\sum_{i\in E_1}\exp\Big(\alpha\varepsilon_i^{(1)}\Big) - \sum_{i\in E_2\cup E_3}\exp\Big(\alpha\varepsilon_i^{(+1)}\Big)\Big) \\ - \frac{1}{2\sigma_{G1}^2 C_{k1}(1-\rho_{12}^2)}\left(\beta_j^{(1)} - \frac{\sigma_{G1}\sqrt{C_{k1}}\rho_{12}\beta_j^{(2)}}{\sigma_{G2}\sqrt{C_{l2}}}\right)^2.\end{aligned} \tag{34}$$

$$\log p(\beta_j^{(2)}|D,\mu,\alpha,(\beta^{(k)})_{k=1,3},(\sigma_{Gk}^2)_{k=1}^3,(\gamma_j^{(k)})_{k=1}^3=(k,l,m)) = -\alpha\beta_j^{(2)}\sum_{i\in E_2} d_i x_{ij}$$
$$+\exp(-K)\Big(\sum_{i\in E_2\cup E_3}\exp\Big(\alpha\varepsilon_i^{(-1)}\Big)-\sum_{i\in E_2}\exp\Big(\alpha\varepsilon_i^{(2)}\Big)-\sum_{i\in E_3}\exp\Big(\alpha\varepsilon_i^{(+2)}\Big)\Big)$$
$$-\frac{1}{2\sigma_{G2}^2 C_{l2}\Big[1-\frac{\rho_{12}^2+\rho_{23}^2-2\rho_{12}^2\rho_{23}^2}{1-\rho_{12}^2\rho_{23}^2}\Big]}\Big(\beta_j^{(2)}-\Big(\frac{\sigma_{G2}\sqrt{C_{l2}}\rho_{23}(1-\rho_{12}^2)\beta_j^{(3)}}{\sigma_{G3}\sqrt{C_{m3}}(1-\rho_{12}^2\rho_{23}^2)}+\frac{\sigma_{G2}\sqrt{C_{l2}}\rho_{12}(1-\rho_{23}^2)\beta_j^{(1)}}{\sigma_{G1}\sqrt{C_{k1}}(1-\rho_{12}^2\rho_{23}^2)}\Big)\Big)^2. \tag{35}$$

$$\log p(\beta_j^{(3)}|D,\mu,\alpha,(\beta^{(k)})_{k=2,3},(\sigma_{Gk}^2)_{k=1}^3,(\gamma_j^{(k)})_{k=1}^3=(k,l,m)) = -\alpha\beta_j^{(3)}\sum_{i\in E_3} d_i x_{ij}$$
$$+\exp(-K)\Big(\sum_{i\in E_3}\Big[\exp\Big(\alpha\varepsilon_i^{(-2)}\Big)-\exp\Big(\alpha\varepsilon_i^{(3)}\Big)\Big]\Big)-\frac{1}{2\sigma_{G3}^2 C_{m3}(1-\rho_{23}^2)}\Big(\beta_j^{(3)}-\frac{\sigma_{G3}\sqrt{C_{m3}}\rho_{23}\beta_j^{(2)}}{\sigma_{G2}\sqrt{C_{l2}}}\Big)^2. \tag{36}$$

$$\log p(\delta_j|D,\mu,\alpha,(\beta^{(k)})_{k=1}^3) = \text{const} - \alpha\sum_{k=1}^3\sum_{i\in E_k} d_i z_{ij}\delta_j - \frac{1}{2\sigma_\delta^2}\delta^2$$
$$+\exp(-K)\sum_{k=1}^3\sum_{i\in E_k}\Big[-\sum_{j=1}^{k-1}\exp\Big(\alpha\varepsilon_i^{(+j)}\Big)-\exp\Big(\alpha\varepsilon_i^{(k)}\Big)+\sum_{j=1}^{k-1}\exp\Big(\alpha\varepsilon_i^{(-j)}\Big)\Big]. \tag{37}$$

We define $\eta_{kl}^{(1)} := \begin{cases}\rho_{12}\frac{\sigma_{G1}\sqrt{C_{k1}}}{\sigma_{G2}\sqrt{C_{l2}}}\beta_j^{(2)}, & l\neq 0\\ 0, & l=0\end{cases}$, $\eta_{ml}^{(3)} := \begin{cases}\frac{\sigma_{G3}\sqrt{C_{m3}}\rho_{23}\beta_j^{(2)}}{\sigma_{G2}\sqrt{C_{l2}}}, & l\neq 0\\ 0, & l=0\end{cases}$ and

$$\eta_{l,km}^{(2)} := \begin{cases}\frac{\sigma_{G2}\sqrt{C_{l2}}}{\sigma_{G1}\sqrt{C_{k1}}}\frac{\rho_{12}(1-\rho_{23})^2}{1-\rho_{12}^2\rho_{23}^2}\beta_j^{(1)}+\frac{\sigma_{G2}\sqrt{C_{l2}}}{\sigma_{G3}\sqrt{C_{m3}}}\frac{\rho_{23}(1-\rho_{12})^2}{1-\rho_{12}^2\rho_{23}^2}\beta_j^{(3)}, & k,m\neq 0\\ \frac{\sigma_{G2}\sqrt{C_{l2}}}{\sigma_{G1}\sqrt{C_{k1}}}\frac{\rho_{12}(1-\rho_{23})^2}{1-\rho_{12}^2\rho_{23}^2}\beta_j^{(1)}, & k\neq 0, m=0\\ \frac{\sigma_{G2}\sqrt{C_{l2}}}{\sigma_{G3}\sqrt{C_{m3}}}\frac{\rho_{23}(1-\rho_{12})^2}{1-\rho_{12}^2\rho_{23}^2}\beta_j^{(3)}, & k=0, m\neq 0\\ 0, & k=m=0\end{cases}.$$

$$\hat{\sigma}_{k1} = \frac{1}{\sqrt{2}\sqrt{1+\alpha^2\sigma_{G1}^2C_{k1}(1-\rho_{12}^2)\Big(\sum_{i\in E_1}x_{ij}^2\exp\Big(u_i^{(+1)}-\alpha x_{ij}\eta_{kl}^{(1)}\Big)+\sum_{i\in E_2\cup E_3}x_{ij}^2\exp\Big(u_i^{(1)}-\alpha x_{ij}\eta_{kl}^{(1)}\Big)\Big)}}.$$

$$\frac{1}{\hat{\sigma}_{l2}^2} = 2\Big(1+\alpha^2\sigma_{G2}^2C_{l2}\Big[1-\frac{\rho_{12}^2+\rho_{23}^2-2\rho_{12}^2\rho_{23}^2}{1-\rho_{12}^2\rho_{23}^2}\Big]\cdot\Big(-\sum_{i\in E_2}x_{ij}^2\exp\Big(u_i^{(2)}-\alpha x_{ij}\eta_{l,km}^{(2)}\Big)$$
$$+\sum_{i\in E_2\cup E_3}x_{ij}^2\exp\Big(u_i^{(-1)}-\alpha x_{ij}\eta_{l,km}^{(2)}\Big)\Big)-\sum_{i\in E_3}x_{ij}^2\exp\Big(u_i^{(+2)}-\alpha x_{ij}\eta_{l,km}^{(2)}\Big)\Big)\Big).$$

$$\frac{1}{\hat{\sigma}_{m3}^2} = 2\Big(1+\alpha^2\sigma_{G3}^2C_{m3}(1-\rho_{23}^2)\cdot\Big(-\sum_{i\in E_3}x_{ij}^2\Big[\exp\Big(u_i^{(-2)}-\alpha x_{ij}\eta_{ml}^{(3)}\Big)-\exp\Big(u_i^{(3)}-\alpha x_{ij}\eta_{ml}^{(3)}\Big)\Big]\Big).$$

$$T_1 = \exp\Bigg(-\sum_{i\in E_1} \exp\Big(u_i^{(1)}\Big) - \sum_{i\in E_2\cup E_3} \exp\Big(u_i^{(+1)}\Big)\Bigg), \quad D_1 = \sqrt{2\sigma_{G1}^2 C_{k1}(1-\rho_{12}^2)}.$$

$$\begin{aligned}\frac{g^{(1)}(\hat\sigma_{k1}\sqrt{2}t_r)}{T_1} &= \exp\Bigg(-\alpha\big(\hat\sigma_{k1}\sqrt{2}t_r D_1 + \eta_{kl}^{(1)}\big)\sum_{i\in E_1} d_i x_{ij} - \sum_{i\in E_1}\exp\Big(u_i^{(1)} - \alpha x_{ij}(\hat\sigma_{k1}\sqrt{2}t_r D_1 + \eta_{kl}^{(1)})\Big) \\ &\quad - \sum_{i\in E_2\cup E_3}\exp\Big(u_i^{(+1)} - \alpha x_{ij}(\hat\sigma_k\sqrt{2}t_r D + \eta_{kl}^{(1)})\Big) + \sum_{i\in E_1}\exp\Big(u_i^{(1)}\Big) + \sum_{i\in E_2\cup E_3}\exp\Big(u_i^{(+1)}\Big) - (\hat\sigma_{k1}\sqrt{2}t_r)^2\Bigg)\end{aligned}$$

$$\begin{aligned}T_2 &= \exp\Bigg(\sum_{i\in E_2\cup E_3}\exp\Big(u_i^{(-1)}\Big) - \sum_{i\in E_2}\exp\Big(u_i^{(2)}\Big) - \sum_{i\in E_3}\exp\Big(u_i^{(+2)}\Big)\Bigg), \\ D_2 &= \sqrt{2\sigma_{G2}^2 C_{l2}\Bigg(1 - \frac{\rho_{12}^2 + \rho_{23}^2 - 2\rho_{12}^2\rho_{23}^2}{1-\rho_{12}^2\rho_{23}^2}\Bigg)}.\end{aligned}$$

$$\begin{aligned}\frac{g^{(2)}(\hat\sigma_{l2}\sqrt{2}t_r)}{T_2} &= \exp\Bigg(-\alpha(D_2\hat\sigma_{l2}\sqrt{2}t_r + \eta_{l;km}^{(2)})\sum_{i\in E_2} d_i x_{ij} \\ &\quad + \sum_{i\in E_2\cup E_3}\Bigg[\exp\Big(u_i^{(-1)} - \alpha x_{ij}(D_2\hat\sigma_{l2}\sqrt{2}t_r + \eta_{l;km}^{(2)})\Big) - \exp\Big(u_i^{(-1)}\Big)\Bigg] \\ &\quad - \sum_{i\in E_2}\Bigg[\exp\Big(u_i^{(2)} - \alpha x_{ij}(D_2\hat\sigma_{l2}\sqrt{2}t_r + \eta_{l;km}^{(2)})\Big) - \exp\Big(u_i^{(2)}\Big)\Bigg] \\ &\quad - \sum_{i\in E_3}\Bigg[\exp\Big(u_i^{(-2)} - \alpha x_{ij}(D_2\hat\sigma_{l2}\sqrt{2}t_r + \eta_{l;km}^{(2)})\Big) - \exp\Big(u_i^{(-2)}\Big)\Bigg] - (\hat\sigma_{l2}\sqrt{2}t_r)^2\Bigg)\end{aligned}$$

$$T_3 = \exp\Bigg(\sum_{i\in E_3}\Bigg[\exp\Big(u_i^{(-2)}\Big) - \exp\Big(u_i^{(3)}\Big)\Bigg]\Bigg), \quad D_3 = \sqrt{2\sigma_{G3}^2 C_{m3}(1-\rho_{23}^2)}.$$

$$\begin{aligned}\frac{g^{(3)}(\hat\sigma_{m3}\sqrt{2}t_r)}{T_3} &= \exp\Bigg[-\alpha(D_3\hat\sigma_{m3}\sqrt{2}t_r + \eta_{ml}^{(3)})\sum_{i\in E_3} d_i x_{ij} \\ &\quad + \sum_{i\in E_3}\Bigg[\exp\Big(u_i^{(-2)} - \alpha x_{ij}(D_3\hat\sigma_{m3}\sqrt{2}t_r + \eta_{ml}^{(3)})\Big) - \exp\Big(u_i^{(-2)}\Big) - \exp\Big(u_i^{(3)} - \alpha x_{ij}(D_3\hat\sigma_{m3}\sqrt{2}t_r + \eta_{ml}^{(3)})\Big) \\ &\quad + \exp\Big(u_i^{(3)}\Big)\Bigg] - (\hat\sigma_{m3}\sqrt{2}t_r)^2\Bigg]\end{aligned}$$

Derivation for the sparse calculations 82

$$\begin{aligned}\hat\sigma_{k1} &= \frac{1}{\sqrt{2}}\Bigg[1 + \alpha^2\sigma_{G1}^2 C_{k1}(1-\rho_{12}^2)\frac{N_j}{s_j^2}\Bigg((\bar\xi_j)^2\Bigg(\sum_{\substack{i\in E_2\cup E_3\\ \xi_{ij}=0}}\exp\Big(u_i^{(1)}\Big) + \sum_{\substack{i\in E_1\\ \xi_{ij}=0}}\exp\Big(u_i^{(+1)}\Big)\Bigg) \\ &\quad + M_j\Bigg((1-\bar\xi)^2\Bigg(\sum_{\substack{i\in E_2\cup E_3\\ \xi_{ij}=1}}\exp\Big(u_i^{(1)}\Big) + \sum_{\substack{i\in E_1\\ \xi_{ij}=1}}\exp\Big(u_i^{(+1)}\Big)\Bigg) + (2-\bar\xi_j)^2\Bigg(\sum_{\substack{i\in E_2\cup E_3\\ \xi_{ij}=2}}\exp\Big(u_i^{(1)}\Big) + \sum_{\substack{i\in E_1\\ \xi_{ij}=2}}\exp\Big(u_i^{(+1)}\Big)\Bigg)\Bigg)\Bigg)\Bigg]^{0.5}.\end{aligned}$$

Expressions for $\hat\sigma_{l2}$ and $\hat\sigma_{m3}$ follow the same pattern - each time there is a sum over an individuals from a certain epoch, we break it into three cases: $\xi_{ij} = 0$, $\xi_{ij} = 1$ and $\xi_{ij} = 2$, we group them together, remove the term of the form $\alpha x_{ij}\eta_{kl}$ by adding the appropriate multiplicative terms in front. We rewrite parts of the equation (34) in the following way:

$$\sum_{i\in E_1}\exp\Big(\alpha\varepsilon_i^{(1)}-K\Big)+\sum_{i\in E_2\cup E_3}\exp\Big(\alpha\varepsilon_i^{(+1)}-K\Big)=G_j\Bigg[\sum_{\substack{i\in E_1\\ \xi_{ij}=0}}\exp\Big(u_i^{(1)}\Big)+\sum_{\substack{i\in E_2\cup E-3\\ \xi_{ij}=0}}\exp\Big(u_i^{(+1)}\Big)$$
$$+H_j\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=1}}\exp\Big(u_i^{(1)}\Big)+\sum_{\substack{i\in E_2\cup E_3\\ \xi_{ij}=1}}\exp\Big(u_i^{(+1)}\Big)+H_j\Bigg(\sum_{\substack{i\in E_1\\ \xi_{ij}=2}}\exp\Big(u_i^{(1)}\Big)+\sum_{\substack{i\in E_2\cup E_3\\ \xi_{ij}=2}}\exp\Big(u_i^{(+1)}\Big)\Bigg)\Bigg)\Bigg].$$

Equations (35) and (36) can be rewritten in an analogous way.

References

1. Joshi, P. K. *et al.* Genome-wide meta-analysis associates HLA-DQA1/DRB1 and LPA and lifestyle factors with human longevity. *Nature Communications* **8**, 910 (2017).

2. Joshi, P. K. *et al.* Variants near CHRNA3/5 and APOE have age- and sex-related effects on human lifespan. *Nature Communications* **7**, 11174 (2016).

3. Tibshirani, R. The lasso method for variable selection in the Cox model. *Statistics in Medicine* **16**, 385–395 (1997).

4. Qian, J. *et al.* A fast and scalable framework for large-scale and ultrahigh-dimensional sparse regression with application to the UK Biobank. *PLOS Genetics* **16**, 1–30 (2020).

5. Li, R. *et al.* Fast Lasso method for large-scale and ultrahigh-dimensional Cox model with applications to UK Biobank. *Biostatistics* (2020).

6. He, L. & Kulminski, A. M. Fast Algorithms for Conducting Large-Scale GWAS of Age-at-Onset Traits Using Cox Mixed-Effects Models. *Genetics* **215**, 41–58 (2020).

7. Ojavee, S. E. *et al.* Genomic architecture and prediction of censored time-to-event phenotypes with a bayesian genome-wide analysis. *Nature Communications* **12**, 2337 (2021). URL `https://doi.org/10.1038/s41467-021-22538-w`.

8. Ge, T., Chen, C.-Y., Neale, B. M., Sabuncu, M. R. & Smoller, J. W. Phenome-wide heritability analysis of the uk biobank. *PLOS Genetics* **13**, 1–21 (2017). URL `https://doi.org/10.1371/journal.pgen.1006711`.

Figures

Supplementary Figures

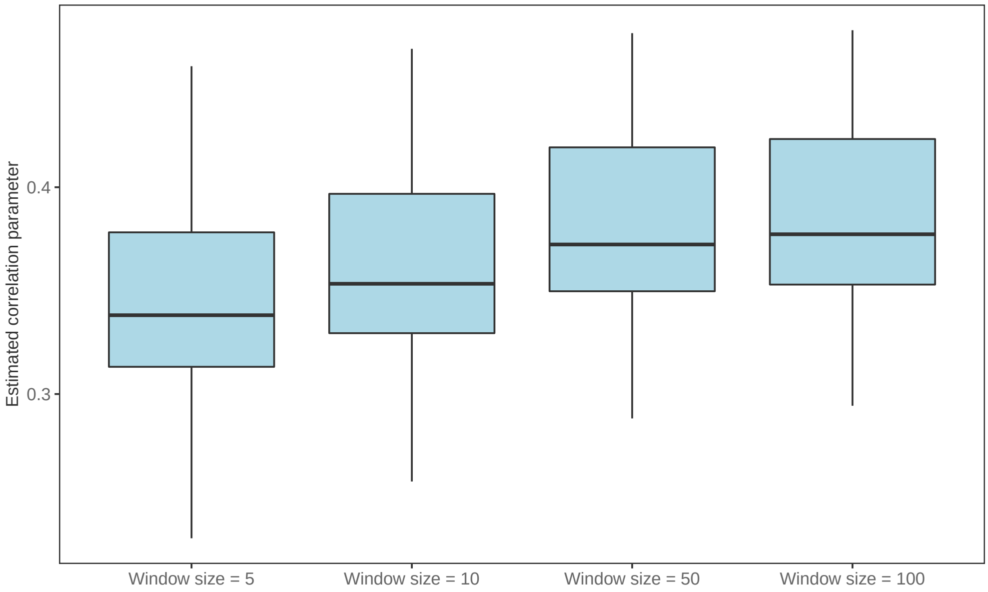

Figure 1. Effect of window size on the inter-epoch correlation. A total of $p = 400$ non-zero effects were generated on top of M =40,249 markers from chromosome 22 using N =8000 randomly selected UK Biobank individuals in 10 independent simulations. The true correlation between epochs was simulated $\rho = 0.5$ and heritability $h^2 = 0.5$. The time-varying BayesW models were estimated with mixture components 0.001,0.01 and by varying window size (the number of last iterations used for estimating genetic correlation). Combining more iterations gives more accurate genetic value estimates yielding more accurate genetic inter-epoch correlation estimates.

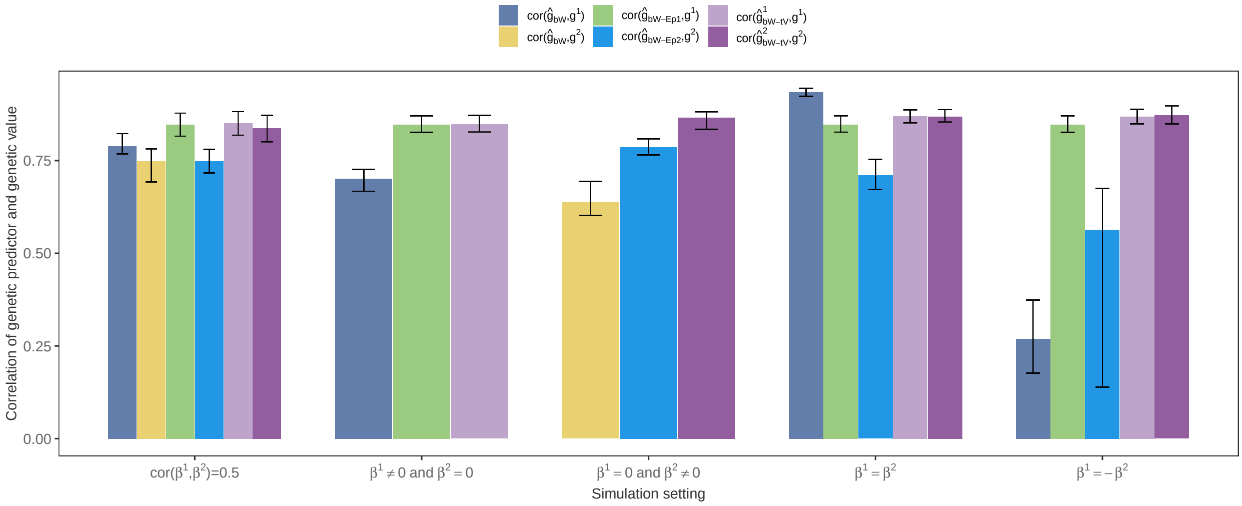

Figure 2. Prediction accuracy when predicting . A total of p = 400 non-zero effects were generated on top of M =40,249 markers from chromosome 22 using N =8000 randomly selected UK Biobank individuals in 10 independent simulations. The assignment of markers between the group was random. The models were estimated using time-varying BayesW model with mixture components 0.001,0.01. In all of the scenarios genetic variances were kept constant (corresponding to heritability h=0.5) within groups A and B and epochs 1 and 2. The inter-epoch correlation was also kept constant in group A but in group B we simulated intra-epoch correlations of 0, 1, -1 and 0.5. Although the groups are random and therefore the markers between groups correlated we manage to recover the group-specific genetic variances and slightly (absolutely) underestimated correlation parameters. The bounds of the box show the interquartile range, centre shows the median and minimum and maximum indicate the 95% credibility interval.

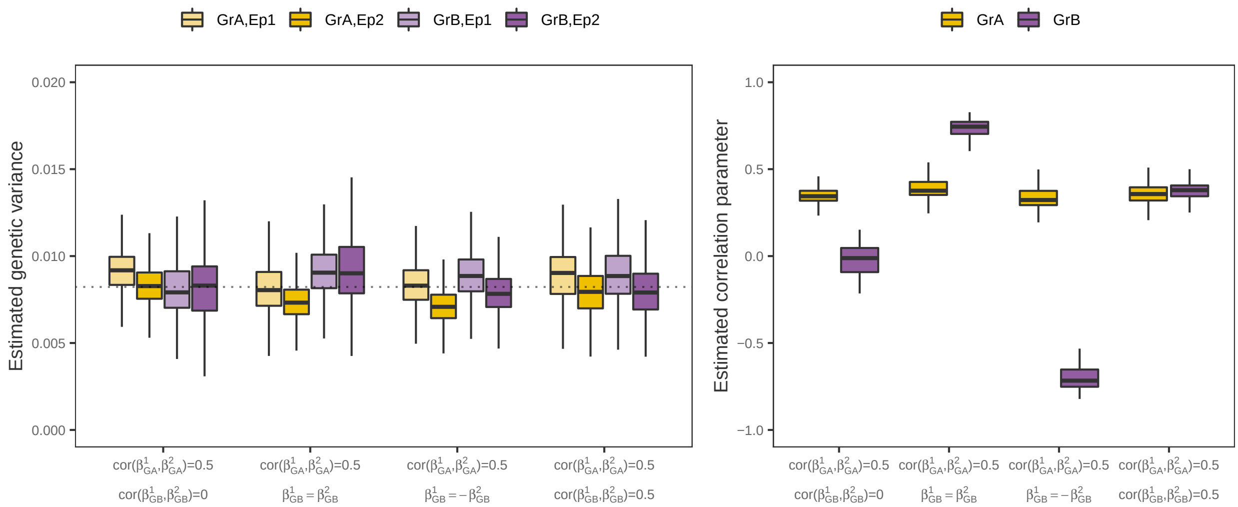

Figure 3. Recovering genetic variance and correlation hyperparameters in case of two groups with constant genetic variances and differing inter-epoch correlations. A total of p = 400 non-zero effects were generated on top of M =40,249 markers from chromosome 22 using N =8000 randomly selected UK Biobank individuals in 10 independent simulations. The assignment of markers between the group was random. The models were estimated using time-varying BayesW model with mixture components 0.001,0.01. In all of the scenarios genetic variances were kept constant (corresponding to heritability h=0.5) within groups A and B and epochs 1 and 2. The inter-epoch correlation was also kept constant in group A but in group B we simulated intra-epoch correlations of 0, 1, -1 and 0.5. Although the groups are random and therefore the markers between groups correlated we manage to recover the group-specific genetic variances and slightly (absolutely) underestimated correlation parameters. The bounds of the box show the interquartile range, centre shows the median and minimum and maximum indicate the 95% credibility interval.

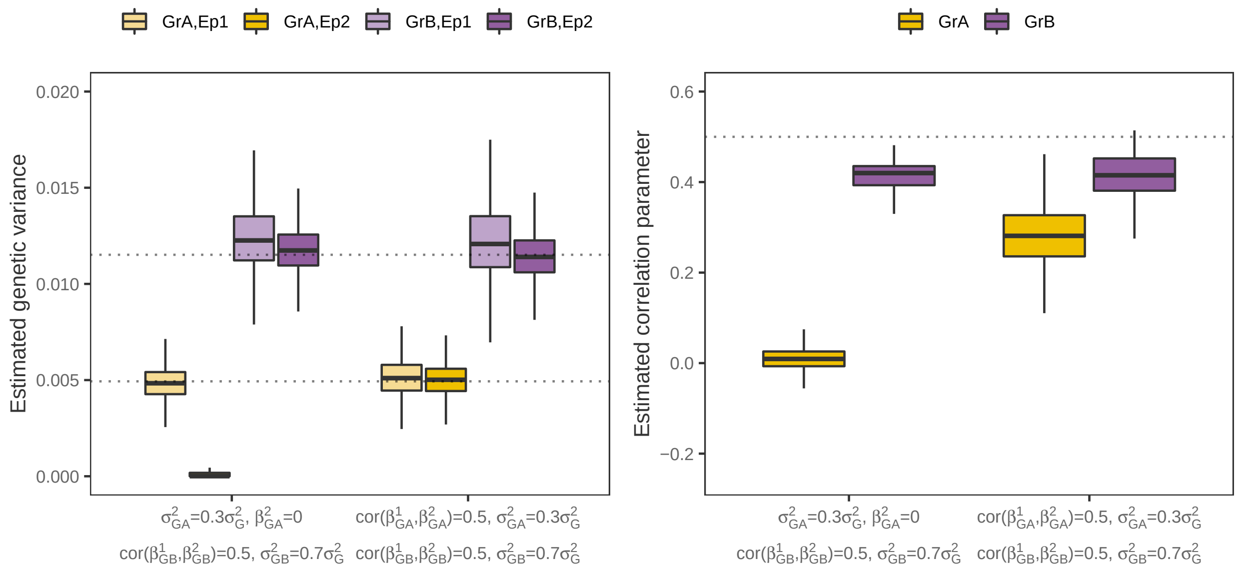

Figure 4. Recovering genetic variance and correlation hyperparameters in case of two groups in two scenarios. A total of $p = 400$ non-zero effects were generated on top of M =40,249 markers from chromosome 22 using N =8000 randomly selected UK Biobank individuals in 10 independent simulations. The assignment of markers between the group was random. The models were estimated using time-varying BayesW model with mixture components 0.001,0.01. In the first scenario, the markers in group A were assigned genetic variance of $0.3\sigma_G^2$ in the first epoch and no genetic variance in the second epoch; in group B markers were assigned genetic variance of $0.7\sigma_G^2$ in both epochs with a correlation of 0.5 between the epochs. In the second scenario, the markers in group A and group B were assigned in both epochs genetic variances of $0.3\sigma_G^2$ and $0.7\sigma_G^2$ respectively with correlations of 0.5 between the epochs. The genetic variance σ_G^2 was chosen such to represent $h^2 = 0.5$. The group-specific genetic variance parameters are captured within the 95% credibility interval but due to imperfect effect estimation, the genetic correlation parameter is slightly underestimated and this effect is stronger in the case of lower genetic variance. The bounds of the box show the interquartile range, centre shows the median and minimum and maximum indicate the 95% credibility interval.

Supplementary Information

Supplementary Note

Generating the phenotypes with different effect sizes

In order to generate data that has an underlying survival function given in equation 4, we use the idea of inverse transform sampling. Based on the specification of survival function in equation 4 it is possible to find the cumulative distribution function and its respective inverse which we then use for sampling. The rule can be summarised as:

Algorithm 1: Generating data from distribution with changing effect sizes.

Data: Intercept μ, shape α, effect size vectors β^1, β^2, marker values x_i for each individual, cut-off point between epochs τ, K is the Euler-Mascheroni constant.

foreach *individual* i **do**
    Generate $u_i \sim U(0,1)$
    if $u_i \geq 1 - \exp\bigl[-\exp\bigl(\alpha(\log(\tau) - \mu - x_i\beta^1) - K\bigr)\bigr]$ **then**
        Generate t_i as:
        $\log(t_i) = \mu + x_i\beta^2 + \frac{K}{\alpha} + \frac{1}{\alpha}\log[-\exp\bigl(\alpha(\log(\tau) - \mu - x_i\beta^1) - K\bigr) + \exp\bigl(\alpha(\log(\tau) - \mu - x_i\beta^2) - K\bigr) - \log(1 - u_i)]$
    else
        Generate t_i as
        $\log(t_i) = \mu + x_i\beta^1 + \frac{K}{\alpha} + \frac{1}{\alpha}\log[-\log(1 - u_i)]$

Moment generating function

Let us rewrite probability density function as

$$f(t) = \begin{cases} \lambda_1 \alpha t^{\alpha-1} e^{-\lambda_1 t^\alpha}, & 0 < t < \tau \\ \lambda_2 \alpha t^{\alpha-1} e^{\tau^\alpha(\lambda_2 - \lambda_1) - \lambda_2 t^\alpha}, & t \geq \tau \end{cases}$$

where $\lambda_k = e^{-\alpha\mu - x_i\beta^k - K}, k \in 1, 2$. Then the probability density function of $Y = \log T$ is

$$g(y) = \begin{cases} \lambda_1 \alpha e^{y(\alpha+t) - \lambda_1 e^{y\alpha}}, & y < \log(\tau) \\ \lambda_2 \alpha e^{y(\alpha+t) - \tau^\alpha(\lambda_2 - \lambda_1) - \lambda_2 e^{y\alpha}}, & y \geq \log(\tau) \end{cases}$$

Then,

$$m(t) = \mathbb{E}[e^{tY}] = \lambda_1^{\frac{-t}{\alpha}} \gamma\left(\frac{\alpha+t}{\alpha}, \lambda_1\tau^\alpha\right) + \lambda_2^{\frac{-t}{\alpha}} e^{-\tau^\alpha(\lambda_2-\lambda_1)} \left(\Gamma\left(\frac{\alpha+t}{\alpha}\right) - \gamma\left(\frac{\alpha+t}{\alpha}, \lambda_2\tau^\alpha\right)\right)$$

$$\frac{\mathrm{d}}{\mathrm{d}s}\gamma(s,x) = \log(x)\Gamma(s,x) + xT(3,s,x)\ldots$$

$$f(t) = \frac{\beta\lambda}{\Gamma(\alpha)}(\lambda t)^{\alpha\beta-1} e^{-(\lambda t)^\beta}, t > 0$$

We fit a Cox proportional-hazards model [**?**] to describe how CBC levels and obstetric complications influence birth rate at a particular point in time. The model is expressed as follow:

$$h_i(t) = h_0(t)exp(\beta_1 \cdot x_i^1 + \beta^2 \cdot x_i^2 + ... + \beta_p \cdot x_i^p) \tag{38}$$

where t represents the pregnancy time, $h_i(t)$ is the hazard function for individual i which can be interpreted as the risk of labour at time t. The regression coefficients $\beta_1, \beta_2, ..., \beta_p$ measure the effect size of covariates. $h_0(t)$ is the baseline hazard function describing how the risk of birth changes over time at baseline covariate levels.