1 Further details of the fitted models for the continuous CES-D HERS data

1.1 Cholesky decomposition for variance components

We use a Cholesky decomposition (Daniels and Zhao, 2003) for modelling the variance components as a function of $u$ in Section 2.1 of the article and in the HERS analysis reported in Section 4.1 of the article.

Recall that the random effects $b_i \sim N(0, G_i)$. Let $\tilde{b}_{ik} (k = 1, \ldots, q)$ be the linear least-squares predictor of the $k$th random effect $b_{ik}$ based on its predecessors $b_{i,k-1}, \ldots, b_{i1}$, and let $e_{ik} = b_{ik} - \tilde{b}_{ik}$ be the prediction error with variance $\sigma^2_{ik} = \text{var}(e_{ik})$; hence

$$b_{ik} = \sum_{s=1}^{k-1} \lambda_{iks} b_{is} + e_{ik}. \quad (1.1)$$

The special Cholesky decomposition of $G_i$ is defined as $L_i G_i L_i' = D_i$, where $L_i$ is the unit lower triangular matrix with $-\lambda_{iks}$ as its $(k, s)$th entry and $D_i = \text{diag}(\sigma^2_{i1}, \ldots, \sigma^2_{iq})$. The $\lambda_{iks}$ are referred to as generalized autoregressive parameters (GARP) and $\sigma^2_{ik}$ as innovation variances (IV). The only constraint needed for $G_i$ to be positive definite is that $\sigma^2_{ik} > 0$ for all $(i,k)$. 

*Email: li.su@mrc-bsu.cam.ac.uk; phone: 44-1223-330386; fax: 44-1223-330388
In the HERS analysis we assume a simple case of random intercept and slope $b_i = (b_{i1}, b_{i2})^T$, and (1.1) can be written in two parts:

$$
\begin{align*}
  b_{i1} &= e_{i1}, \\
  b_{i2} &= \lambda_{i21} b_{i1} + e_{i2},
\end{align*}
$$

where $\text{var}(e_{ik}) = \sigma_{ik}^2$, $k = 1, 2$. The first equation corresponds to the marginal distribution of the random intercepts, and the second equation describes the conditional distribution of random slopes given random intercepts. For the LMM and VCM1 in Section 4.1 of the article, we use this parametrization for variance components, but do not allow them to vary by the dropout/administrative censoring time. For VCM2, given a dropout/administrative censoring time $u_i$, the GARP and IV parameters are modeled as follows:

$$
\begin{align*}
  \lambda_{i21} | u_i, \delta_i &= \lambda_{21, \delta_i}(u_i) \\
  \{\log(\sigma_{ik}^2) | u_i, \delta_i\} &= \sigma_{k, \delta_i}^2(u_i), \quad k = 1, 2.
\end{align*}
$$

Further, the error variance matrix $R_i$ is also parameterized such that it depends on $u$. We assume that the errors are independent over time, $R_i = \tau_i^2 I_i$, and

$$
\{\log(\tau_i^2) | u_i, \delta_i\} = h_{\delta_i}(u_i).
$$

Therefore, $\theta_{\delta_i}(u) = \{\beta_{\delta_i}(u)^T, \lambda_{21, \delta_i}(u), \sigma_{1, \delta_i}^2(u), \sigma_{2, \delta_i}^2(u), h_{\delta_i}(u)\}^T$ in VCM2 for the continuous HERS data. In this application because the administrative censoring times are similar, we then assume that for the administrative censoring group, $\theta_0(u)$ are constants, while for the dropout group, $\theta_1(u)$ are smooth functions modelled by Bayesian penalized splines. Other forms for $\theta_0(u)$ can be assumed in applications where individuals in the administrative censoring group are heterogeneous.

### 1.2 Prior specification and posterior inference

For all smooth functions, the prior specification described in Section 3 of the article is used for the parameters in the 9-knot (placed at the deciles of the observed dropout times) low-rank thin-plate basis penalized splines. In the LMM and the VCMs, Normal priors with
zero mean and large variances are used for the fixed effects $\beta$ and $\beta_0$, respectively. For both the LMM and VCM1, Normal priors with zero mean and large variances are assigned to the variance components in $G$ and $\tau^2$ with the Cholesky decomposition.

We use Cox regression analysis to check the relationship between the covariates and the dropout time. The Whites and Blacks were less likely to drop out than the Latinas (including others); the patients with baseline CD4 > 200 were also less likely to drop out. Therefore, we have $f(u, \delta \mid x) \neq f(u, \delta)$. For summarizing marginal covariates effects in the VCMs, Bayesian bootstrapping needs to be conducted within covariates. Because the dropout/administrative censoring time distribution depends on the race and baseline CD4 groups, we are not able to provide simple summaries of the marginal covariate effects in the VCMs. However, it can be shown that the estimated CES-D profiles within the race and baseline CD4 groups are still linear.

We run two MCMC chains with diverse initial values and assess convergence within a 5,000-iteration burn-in period using history plots and Gelman and Rubin convergence statistics provided by the WinBUGS package. After convergence, pooled posterior samples of size 20,000 are used for model inference.

2 Additional results for the continuous CES-D HERS data

Figure 1 plots the observed CES-D data. Histograms and Normal Quantile-Quantile plots show that the posterior means for the residuals in all models are approximately Normal. Figure 2, Tables 1 and 2 give the additional results from the VCMs. The posterior mean estimates for $\beta_1(u)$ in the dropout group and $\beta_0$ in the administrative censoring group follow similar patterns in both VCMs.
2.1 Model assessment and selection

Model assessment and selection is an open area of research for handling informative dropout. None of the existing diagnostics and metrics can evaluate the feasibility of the assumptions about the missing data given the observed data (Dobson and Henderson, 2003; Daniels and Hogan, 2008). Instead, they can only provide information on whether joint modelling assumptions, priors and assumed missing data mechanisms are compatible with the observed data. In fact, Molenberghs and others (2008) have shown that every MNAR model has a MAR counterpart with equal fit. For instance, in our HERS example, the values of Deviance Information Criterion (DIC) (Spiegelhalter and others, 2002) based on the observed data from the LMM and the two VCMs are 46807.4, 46810.7 and 46814.1, respectively, which shows that these models had similar fit to the observed data.

2.2 Example of sensitivity analysis

Sensitivity analysis is needed to assess the unverifiable assumption implied in the VCM. In the HERS application, it is assumed that the rate of change in CES-D scores beyond \( u \) is the same as the one before \( u \) for individuals who dropped out at \( u \). Here we demonstrate an example of sensitivity analysis regarding this assumption. Basically, we assume a different CES-D slope when \( t > u \), i.e., assume a continuous piece-wise linear model with a change point at \( u \). Specifically,

\[
E(Y_{ij} \mid b_i, U_i = u, \delta_i = 1) = \beta_0(u) + \beta_1(u) \cdot I(\text{Black}) + \beta_2(u) \cdot I(\text{White}) \\
+ \beta_3(u) \cdot I(\text{CD4} > 200) + \beta_4(u_i) \cdot t_{ij} \\
+ \beta_5(u) \cdot I(\text{CD4} > 200) \cdot t_{ij} + b_{i1} + b_{i2} \cdot t_{ij} + \omega(u)(t_{ij} - u)_+, 
\]

where \((x)_+ = x\) if \(x > 0\) and 0 otherwise, and \(\omega(u)\) is the change of the CES-D slope beyond \(u\). The parameterization for the variance components is the same as in Section 4.1 of the article.

In principle, sensitivity analysis should be based on the parameters that cannot be identified by the observed data, such as \(\omega(u)\). We assume a simple functional form,
\[ \omega(u) = a\{\max(U) - u\}/\text{range}(U). \] Note that we do not adjust the CES-D slopes for the administrative censoring group.

We can fix \( a \) at various values, and recompute quantities of interest (such as expected CES-D profiles) to check their sensitivity to \( a \). Figure 3 presents the estimated CES-D profiles when \( a = 5 \) and \( a = 10 \). The estimates at early study period are close across three models. All CES-D trends are adjusted upward further at the later study period when \( a = 5 \) and \( a = 10 \) compared with the estimates from VCM2. In practice, we could specify a range or informative priors for the sensitivity parameters based on expert opinions and prior elicitation from previous studies (Lee, 2007).

[Figure 3 about here.]

3 Further details of the fitted models for the binary CES-D HERS data

In preliminary analysis using varying coefficient MTM(1), for the 580 patients who dropped out before finishing 12 scheduled visits, we let the covariate effects \( \beta_1(u) \) and the serial dependence \( \alpha(u) \) be smooth functions modelled by penalized splines with 9-knot low-rank thin-plate bases (knots are placed at the deciles of the observed dropout times). For the 173 patients with administratively censored dropout times, separate parameters \( \beta_0 \) and \( \alpha_0 \) are assumed. The same prior specification for penalized splines is used as in Section 3 of the article and we assign vague Normal priors to \( \beta_0 \) and \( \alpha_0 \). For \( \beta \) and \( \alpha \) in the original MTM(1), vague Normal priors with mean zero are assigned.

Preliminary results showed that the intercept and the race effects do not have obvious patterns over the observed dropout times. For illustration and for simplicity, we reduce our varying coefficient MTM(1) by only allowing the following parameters to vary by the observed dropout time: the coefficient of baseline CD4 indicator, the coefficient of the time variable, the interaction between time and baseline CD4 count, and the serial dependence parameter. To examine the marginal probability of depression, we apply the method described in Section 3.4 of the article; Bayesian bootstrapping for the observed
dropout/administrative censoring times is conducted within the race and baseline CD4 groups.

For both models, we run two MCMC chains and check the convergence after 5,000-iteration burn-in period using the facilities provided by the WinBUGS package. Pooled posterior samples of size 20,000 are used for inference.

4 Additional results for the binary CES-D HERS data

In the VCM, the posterior mean estimates as well as the 95% credible intervals for the intercept term, the coefficients for the race indicators (Black and White), the baseline CD4 indicator, the time, the interaction between time and baseline CD4 count and the serial dependence in the administrative censoring group are 0.72 [−0.27, 1.64], −0.86 [−1.31, −0.41], −0.63 [−1.11, −0.15], 0.29 [−0.54, 1.19], −0.59 [−1.29, 0.09], 0.42 [−0.30, 1.11] and 2.47 [2.25, 2.69], respectively. The estimated intercept term and coefficients of race indicators in (4.1) for the dropout group are 0.48 [−0.01, 0.95], −0.41 [−0.68, −0.15] and −0.18 [−0.51, 0.14].

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]
5 WinBUGS program for the continuous CES-D HERS data

5.1 Data

Data are sorted by the race and baseline CD4 groups, and consist of the CES-D scores \(y\), the time since enrollment \(t\), the individual indicator \(sub\), the race group indicator \(race\), the indicator of baseline CD4 > 200 \(cd4\), the observed dropout time \(U\), the design matrix of the random effects \(BZ\) for penalized splines, the sample size \(n\), the total number of observations \(Nobs\), the number of observations within individuals \(nobs\), the individual indexes for the race and baseline CD4 groups \(n1,n2\), the knots \(knotD\) and the square root of penalty matrix \(OMEGAu\) for the random effects in low-rank thin-plate penalized splines, the maximum of the dropout times \(MaxU\), the range of the dropout times \(RangeU\), and the sensitivity parameter \(a\).

5.2 Initial values

Initial values are provided for the fixed effects \(a0,a1,a2,a3,a5,Ea2,Ea3,Ea4,Ea5\), the random effects \(d0,d1,d2,d3,d5\), the smoothing parameters \(taud0,taud1,taud2,taud31,taud32,tau5,tau6,tauEd2,tauEd3,tauEd4,tauEd5\). Both data and initial values are specified and processed in R and then used in WinBUGS.

5.3 WinBUGS code

```winbugs
model { # begin model

  for (k in 1:Nobs)
    { y[k]~dnorm(muy[k],taue[sub[k]])
    muy[k]<-beta0[sub[k]]+beta1[sub[k]]*t[k]
    }

  for (i in 1:n)
    {
      beta0[i]<-beta0pop[i]+b[i,1]
      beta0pop[i]<-int[i]+black[i]*equals(race[i],2)
      +white[i]*equals(race[i],3)+cd4coef[i]*cd4[i]
    }

}  # end model
```


9-knot (at sample quantiles (10%) of U[i]) thin plate bases
for smooth functions in the mean structure

# intercept, race and baseline CD4 effects
int[i]<-(a0[1]+a0[2]*U[i]+inprod2(d0[1:9],BZ[i,1:9]))
   *step(11-nobs[i])+intlast*equals(nobs[i],12)

black[i]<-(a1[1]+a1[2]*U[i]+inprod2(d1[1:9],BZ[i,1:9]))
   *step(11-nobs[i])+blacklast*equals(nobs[i],12)

white[i]<-(a2[1]+a2[2]*U[i]+inprod2(d2[1:9],BZ[i,1:9]))
   *step(11-nobs[i])+whitelast*equals(nobs[i],12)

cd4coef[i]<-(a5[1]+a5[2]*U[i]+inprod2(d5[1:9],BZ[i,1:9]))
   *step(11-nobs[i])+cd4coeflast*equals(nobs[i],12)

cd4g[i]<-cd4[i]+1 #baseline CD4 group indicator
betai[i]<-beta1pop[i]+b[i,2]
beta1pop[i]<-(a3[cd4g[i],1]+a3[cd4g[i],2]*U[i]
   +inprod2(d3[cd4g[i],1:9],BZ[i,1:9]))*step(11-nobs[i])
   +beta1last[cd4g[i]]*equals(nobs[i],12)

# modified Cholesky decomposition of random effect covariance matrix
# phi: correlation between random effects
# e: prediction error of random slopes on random intercepts
b[i,1]~dnorm(0, taub0[i])
b[i,2]<-phi[i]*b[i,1]+e[i]
e[i]~dnorm(0, taub1[i])

# 9-knot (at sample quantiles (10%) of U[i]) thin plate bases
for smooth functions in variance components

# error variance
errorsd[i]<-pow(taue[i],-1/2)
taue[i]<-1/exp(sigmae[i])
sigmae[i]<-(Ea2[1]+Ea2[2]*U[i]+inprod2(Ed2[1:9]*BZ[i,1:9]))
   *step(11-nobs[i])+sigmaelast*equals(nobs[i],12)

# variance of random intercepts
sigmab0[i]<-pow(taub0[i],-1/2)
taub0[i]<-1/exp(sigma2b0[i])
sigma2b0[i]<-(Ea3[1]+Ea3[2]*U[i]+inprod2(Ed3[1:9]*BZ[i,1:9]))
   *step(11-nobs[i])+sigma2b0last*equals(nobs[i],12)

# prediction variance of random slopes on random intercepts
sigmab1[i]<-pow(taub1[i],-1/2)
taub1[i]<-1/exp(sigma2b1[i])
sigma2b1[i]<-(Ea4[1]+Ea4[2]*U[i]+inprod2(Ed4[1:9]*BZ[i,1:9]))
   *step(11-nobs[i])+sigma2b1last*equals(nobs[i],12)

# regression coefficient of random slope on random intercepts
phi[i]<-(Ea5[1]+Ea5[2]*U[i]+inprod2(Ed5[1:9]*BZ[i,1:9]))
   *step(11-nobs[i])+philast*equals(nobs[i],12)
# variance components in the administrative censoring group

```r
errorvarlast <- exp(sigmaelast)
varb0last <- exp(sigma2b0last)
varb1last <- exp(sigma2b1last) + exp(sigma2b0last) * pow(philast, 2)
sderrorlast <- sqrt(errorvarlast)
sdb0last <- sqrt(varb0last)
sdb1last <- sqrt(varb1last)
```

# Priors for d0, d1, d2, d3, d5, Ed2-Ed5 (random effects)

```r
for (k in 1:9) {
  d0[k] ~ dnorm(0, tau0)
d1[k] ~ dnorm(0, tau1)
d2[k] ~ dnorm(0, tau2)
d3[1, k] ~ dnorm(0, tau31)
d3[2, k] ~ dnorm(0, tau32)
d5[k] ~ dnorm(0, tau5)
  Ed2[k] ~ dnorm(0, tauEd2)
  Ed3[k] ~ dnorm(0, tauEd3)
  Ed4[k] ~ dnorm(0, tauEd4)
  Ed5[k] ~ dnorm(0, tauEd5)
}
```

# Prior for the smoothing parameters

```r
tau0 ~ dgamma(1.0E-2, 1.0E-2)
tau1 ~ dgamma(1.0E-2, 1.0E-2)
tau2 ~ dgamma(1.0E-2, 1.0E-2)
tau31 ~ dgamma(1.0E-2, 1.0E-2)
tau32 ~ dgamma(1.0E-2, 1.0E-2)
tau5 ~ dgamma(1.0E-2, 1.0E-2)
tauEd2 ~ dgamma(1.0E-2, 1.0E-2)
tauEd3 ~ dgamma(1.0E-2, 1.0E-2)
tauEd4 ~ dgamma(1.0E-2, 1.0E-2)
tauEd5 ~ dgamma(1.0E-2, 1.0E-2)
```

# Priors for the fixed effects

```r
for (k in 1:2) {
  a0[k] ~ dnorm(0, 1.0E-4)
a1[k] ~ dnorm(0, 1.0E-4)
a2[k] ~ dnorm(0, 1.0E-4)
a3[k, 1] ~ dnorm(0, 1.0E-4)
a3[k, 2] ~ dnorm(0, 1.0E-4)
a5[k] ~ dnorm(0, 1.0E-4)
  Ea2[k] ~ dnorm(0, 1.0E-4)
  Ea3[k] ~ dnorm(0, 1.0E-4)
  Ea4[k] ~ dnorm(0, 1.0E-4)
  Ea5[k] ~ dnorm(0, 1.0E-4)
}
```

```r
intlast ~ dnorm(0, 1.0E-4)
blacklast ~ dnorm(0, 1.0E-4)
whitelast ~ dnorm(0, 1.0E-4)
```
cd4coeflast ~ dnorm(0, 1.0E-4)
beta1last[1] ~ dnorm(0, 1.0E-4)
beta1last[2] ~ dnorm(0, 1.0E-4)
interlast <- beta1last[2] - beta1last[1]
sigaelast ~ dnorm(0, 1.0E-4)
sigma2b0last ~ dnorm(0, 1.0E-4)
sigma2b1last ~ dnorm(0, 1.0E-4)
philast ~ dnorm(0, 1.0E-4)

# smooth function estimates
# on 106 grid points in the range of the observed dropout times

for (i in 1:106)
{
  tt[i] <- (i-1)*0.02
  intcur[i] <- Mv(a0[1:2], d0[1:9], knotD[], OMEGAu[], tt[i])
  blackcur[i] <- Mv(a1[1:2], d1[1:9], knotD[], OMEGAu[], tt[i])
  whitecur[i] <- Mv(a2[1:2], d2[1:9], knotD[], OMEGAu[], tt[i])
  cd4coefcur[i] <- Mv(a5[1:2], d5[1:9], knotD[], OMEGAu[], tt[i])
  betac0[i] <- Mv(a3[1,1:2], d3[1,1:9], knotD[], OMEGAu[], tt[i])
  betac1[i] <- Mv(a3[2,1:2], d3[2,1:9], knotD[], OMEGAu[], tt[i])
  inter[i] <- betac1[i] - betac0[i]
  sigmaecur[i] <- Mv(Ea2[1:2], Ed2[1:9], knotD[], OMEGAu[], tt[i])
  sigmab0cur[i] <- Mv(Ea3[1:2], Ed3[1:9], knotD[], OMEGAu[], tt[i])
  sigmab1cur[i] <- Mv(Ea4[1:2], Ed4[1:9], knotD[], OMEGAu[], tt[i])
  phicur[i] <- Mv(Ea5[1:2], Ed5[1:9], knotD[], OMEGAu[], tt[i])

  errorvarcur[i] <- exp(sigmaecur[i])
  b0varcur[i] <- exp(sigmab0cur[i])
  b1varcur[i] <- exp(sigmab1cur[i]) + exp(sigmab0cur[i]) * pow(philast[i], 2)
  sderror[i] <- sqrt(errorvarcur[i])
  sdb0[i] <- sqrt(b0varcur[i])
  sdb1[i] <- sqrt(b1varcur[i])
}

# generate posterior probability for the dropout/adminstrative
censoring times within the race and baseline CD4 groups

for (i in 1:n){del[i] ~ dgamma(1,1)}

for (k in 1:6)
{for (i in nn1[k]:nn2[k])
  { pp[i] <- del[i]/sum(del[nn1[k]:nn2[k]])
    beta0popavg[k] <- inprod2(beta0pop[nn1[k]:nn2[k]], pp[nn1[k]:nn2[k]])
    beta1popavg[k] <- inprod2(beta1pop[nn1[k]:nn2[k]], pp[nn1[k]:nn2[k]])
  }
}
# sensitivity analysis, fix non-identifiable parameter a=5,10
# recompute the marginal CES-D profiles on 106 grid points

for (i in 1:106) {
  tt[i]<-(i-1)*0.02
  for (j in 1:n) {
    omega[j]<-(MaxU-U[j])*a/RangeU
  }
  mar0[i,j]<-pp[j]*(beta0pop[j]+beta1pop[j]*tt[i]
                    +omega[j]*(tt[i]-U[j])*step(tt[i]-U[j]))*step(11-nobs[j])
                    +pp[j]*(beta0pop[j]+beta1pop[j]*tt[i])*equals(nobs[j],12)
  }
  mean0[i]<-sum(mar0[i,nn1[1]:nn2[1]]) # Latina, CD4<=200
  mean1[i]<-sum(mar0[i,nn1[2]:nn2[2]]) # Black, CD4<=200
  mean2[i]<-sum(mar0[i,nn1[3]:nn2[3]]) # White, CD4<=200
  mean01[i]<-sum(mar0[i,nn1[4]:nn2[4]]) # Latina, CD4>200
  mean11[i]<-sum(mar0[i,nn1[5]:nn2[5]]) # Black, CD4>200
  mean21[i]<-sum(mar0[i,nn1[6]:nn2[6]]) # White, CD4>200
}

# end model

6 WinBUGS program for the binary CES-D HERS data

6.1 Data

Data are sorted by the race and baseline CD4 groups, and consist of the binary outcome variable (y[,]), the time since enrollment (t[]), the race group indicator (race[]), the indicator of baseline CD4 > 200 (cd4[]), the observed dropout time (U[]), and the design matrix of the random effects (BZ[,]) for penalized splines, the sample size (n), the number of individuals with only one observation (n1), the number of observations within individuals (nobs[]), the cumulative sum of the observations by individuals (summobs[]), the individual indexes for the race and baseline CD4 groups (n1[],n2[],n3[],n4[]), the knots (knotD[]) and the square root of penalty matrix (OMEGAu[,]) for the random effects in low-rank thin-plate penalized splines.
6.2 Initial values

Initial values are provided for the fixed effects (intdrop, beta1drop, beta2drop, intlast, beta1last, beta2last, beta3last, beta1alllast[], corrlast, a1[], a2[], a3[]), the random effects (d1[], d2[], d3[]), the smoothing parameters (taud1, taud2[], taud3). Both data and initial values are specified and processed in R and then used in WinBUGS.

6.3 WinBUGS code

model {  # begin model
    # data for individuals with only one observation
    for (i in 1:n1)
    {
        # the likelihood of the first observation of each sequence
        # only depends on the marginal mean
        y[i,1]~dbern(p[i,1])
        logit(p[i,1])<-etaM[i,1]
        etaM[i,1]<-beta0all[i]+beta1all[i]*t[(sumn[i]-nobs[i]+1)]

        # data for individuals with more than one observation
        for (i in (n1+1):n)
        {
            y[i,1]~dbern(p[i,1])
            logit(p[i,1])<-etaM[i,1]
            etaM[i,1]<-beta0all[i]+beta1all[i]*t[(sumn[i]-nobs[i]+1)]

            # the likelihood of the rest observations of each sequence
            # depends on both the marginal mean and serial dependence
            for (j in 2:nobs[i])
            {
                y[i,j]~dbern(p[i,j])
                logit(p[i,j])<-delta[i,j]+gamma[i,j]*y[i,j-1]
                etaM[i,j]<-beta0all[i]+beta1all[i]*t[(sumn[i]-nobs[i]+j)]

                # MTM1Delta is a function written in the WinBUGS development
                # interface for computing the intercept delta[i,j] to satisfy
                # marginal mean and serial dependence simultaneously
                delta[i,j]<-MTM1Delta(etaM[i,j],etaM[i,j-1],gamma[i,j])
                gamma[i,j]<-cor[i]
            }
        }
    }

    # 9-knot (at sample quantiles (10%) of U[i]) thin plate bases
    # for smooth functions in regression parameters
    for (i in 1:n)
    {
        # intercept, race and baseline CD4 effects
        beta0all[i]<-int[i]+beta1[i]*equals(race[i],2)+beta2[i]*equals(race[i],3)+beta3[i]*cd4[i]
    }
}
int[i]<-intdrop*step(11-nobs[i])+intlast*equals(nobs[i],12)
beta1[i]<-beta1drop*step(11-nobs[i])+beta1last*equals(nobs[i],12)
beta2[i]<-beta2drop*step(11-nobs[i])+beta2last*equals(nobs[i],12)

beta3[i]<-(a1[1]+a1[2]*U[i]+inprod2(d1[1:9],BZ[i,1:9]))
  *step(11-nobs[i])+beta3last*equals(nobs[i],12)

# time slopes by CD4 groups
cd4g[i]<-cd4[i]+1 #baseline CD4 group indicator

beta1all[i]<-(a2[cd4g[i],1]+a2[cd4g[i],2]*U[i]
  +inprod2(d2[cd4g[i],1:9],BZ[i,1:9]))
  *step(11-nobs[i])+beta1alllast[cd4g[i]]
  *equals(nobs[i],12)

# serial dependence
cor[i]<-(a3[1]+a3[2]*U[i]+inprod2(d3[1:9],BZ[i,1:9]))
  *step(11-nobs[i])+corrlast*equals(nobs[i],12)
}

# Priors for intercept, race effects in the dropout group
intdrop~dnorm(0,1.0E-3)
beta1drop~dnorm(0,1.0E-3)
beta2drop~dnorm(0,1.0E-3)

# Priors for parameters in the administrative censoring group
intlast~dnorm(0,1.0E-3)
beta1last~dnorm(0,1.0E-3)
beta2last~dnorm(0,1.0E-3)
beta3last~dnorm(0,1.0E-3)
for(j in 1:2){beta1alllast[j]~dnorm(0,1.0E-3)}
corrlast~dnorm(0,1.0E-3)
interactionlast<-beta1alllast[2]-beta1alllast[1]

# Priors for d0, d1, d2, d3 (random effects)
for (k in 1:9)
  {
    d1[k]~dnorm(0,taud1)
    for(j in 1:2){d2[j,k]~dnorm(0,taud2[j])}
    d3[k]~dnorm(0,taud3)
  }

# Priors for a0, a1,a2, a3 (fixed effects)
for (k in 1:2)
  {
    a1[k]~dnorm(0,1.0E-3)
    for(j in 1:2){a2[j,k]~dnorm(0,1.0E-3)}
    a3[k]~dnorm(0,1.0E-3)
  }

# Priors for the smoothing parameters
taud1~dgamma(1.0E-2,1.0E-2)
for(j in 1:2){taud2[j]~dgamma(1.0E-2,1.0E-2)}
taud3~dgamma(1.0E-2,1.0E-2)
# smooth function estimates of regression coefficients
# on 106 grid points in the range of the observed dropout times
for (i in 1:106)
{
    v[i]<-(i-1)*0.02
    
    # Mv is a function written in the WinBUGS development
    # interface for computing smooth function estimates
    # with low-rank thin-plate bases
    betacd4[i]<-Mv(a1[1:2], d1[1:9], knotD[], OMEGAu[], v[i])
betc0[i]<-Mv(a2[1,1:2], d2[1,1:9], knotD[], OMEGAu[], v[i])
betc1[i]<-Mv(a2[2,1:2], d2[2,1:9], knotD[], OMEGAu[], v[i])
interaction[i]<-betc1[i]-betc0[i]
corr[i]<-Mv(a3[1:2], d3[1:9], knotD[], OMEGAu[], v[i])
}

# generate posterior probability for the dropout/administrative
# censoring times within the race and baseline CD4 groups
for (i in 1:n)
{
    del[i]~dgamma(1,1)
}

for (k in 1:6)
{
    for (i in nn1[k]:nn2[k])
    {
        pp[i]<-del[i]/(sum(del[nn1[k]:nn2[k]])+sum(del[nn3[k]:nn4[k]]))
    }
    for (i in nn3[k]:nn4[k])
    {
        pp[i]<-del[i]/(sum(del[nn1[k]:nn2[k]])+sum(del[nn3[k]:nn4[k]]))
    }
}

# compute marginal probability estimates averaging over
# Bayesian bootstrap samples, on 106 grid time points
for (j in 1:106)
{
    vv[j]<-(j-1)*0.02
    for (i in 1:n)
    {
        margin[i,j]<-pp[i]*(exp(beta0all[i]+beta1all[i]*vv[j])
                        /(1+exp(beta0all[i]+beta1all[i]*vv[j])))
    }

    # cd4<=200, race=latina&other
    mean0[j]<-(sum(margin[nn1[1]:nn2[1],j])
               +sum(margin[nn3[1]:nn4[1],j]))
    # cd4<=200, race=black
    mean1[j]<-(sum(margin[nn1[2]:nn2[2],j])
               +sum(margin[nn3[2]:nn4[2],j]))
    # cd4<=200, race=white
    mean2[j]<-(sum(margin[nn1[3]:nn2[3],j])
               +sum(margin[nn3[3]:nn4[3],j]))

}
# cd4>200, race=latina&other
mean01[j]<- (sum(margin[nn1[4]:nn2[4],j])
+ sum(margin[nn3[4]:nn4[4],j]))

# cd4>200, race=black
mean11[j]<- (sum(margin[nn1[5]:nn2[5],j])
+ sum(margin[nn3[5]:nn4[5],j]))

# cd4>200, race=white
mean21[j]<- (sum(margin[nn1[6]:nn2[6],j])
+ sum(margin[nn3[6]:nn4[6],j]))

### fix race, compare baseline cd4 groups
latinadiff[j]<-mean01[j]-mean0[j]
blackdiff[j]<-mean11[j]-mean1[j]
whitediff[j]<-mean21[j]-mean2[j]

### fix baseline cd4, compare race groups
cd40BW[j]<-mean1[j]-mean2[j]
cd40LB[j]<-mean0[j]-mean1[j]
cd40LW[j]<-mean0[j]-mean2[j]

cd41BW[j]<-mean11[j]-mean21[j]
cd41LB[j]<-mean01[j]-mean11[j]
cd41LW[j]<-mean01[j]-mean21[j]

}{
}#end model

References

Daniels, M. and Hogan, J. (2008) Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis., vol. 101 of Monographs on Statistics and Applied Probability. New York: Chapman & Hall/CRC.

Daniels, M. and Zhao, Y. (2003) Modelling the random effects covariance matrix in longitudinal data. Statistics in Medicine, 22, 1631–1647.

Dobson, A. and Henderson, R. (2003) Diagnostics for joint longitudinal and dropout time modeling. Biometrics, 59, 741–751.

Lee, J. Y. (2007) Sensitivity Analysis and Informative Priors for Longitudinal Binary Data with Dropout. Ph.D. thesis, Brown University.

Molenberghs, G., Beunckens, C. and Sotto, C. (2008) Every missingness not at random model has a missingness at ranom counterpart with equal fit. Journal of the Royal Statistical Society. Series B, 70, 371–388.

Spiegelhalter, D., Best, N., Carlin, B. and van der Linde, A. (2002) Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society. Series B, 64, 583–639.
Figure 1: The observed continuous CES-D data in the HERS, with 5 individual profiles highlighted.
Figure 2: Estimated smooth functions of the observed dropout times in the mean structure from VCM1 for the continuous CES-D data in the HERS; gray shades are the point-wise 95% credible bands; dashed lines are corresponding estimates. For the 580 patients who dropped out before finishing 12 scheduled visits, we let the covariate effects $\beta_1(u)$ and the serial dependence $\alpha(u)$ be smooth functions modelled by penalized splines with 9-knot low-rank thin-plate bases (knots are placed at the deciles of the observed dropout times). The same prior specification for penalized splines is used as in Section 3 of the article.
Figure 3: Sensitivity analysis for VCM2 for the continuous CES-D data in the HERS: the expected CES-D profiles (posterior means) when the sensitivity parameter $a = 5$ and $a = 10$, compared with the results under the LMM and VCM2; solid line: estimates from VCM2 with $a = 5$; dotdash line: estimates from VCM2 with $a = 10$; dashed line: estimates from VCM2; dotted line: estimates from the LMM.
Figure 4: Marginal probability profiles of depression by the race and baseline CD4 groups for the binary CES-D data in the HERS; left panels: baseline CD4 ≤ 200, right panels: baseline CD4 > 200; solid lines are estimates from VCM MTM(1), dotted lines are estimates from MTM(1) assuming MAR.
Figure 5: Estimated effect of the baseline CD4 count on marginal probability profiles of depression, stratified by race from the VCM for the binary CES-D data in the HERS; gray shades are point-wise 95% credible bands.
Figure 6: The estimated race effects on marginal probability profiles of depression, stratified by baseline CD4 groups from the VCM for the binary CES-D data in the HERS; left panels: baseline CD4 ≤ 200, right panels: baseline CD4 > 200; gray shades are point-wise 95% credible bands.
Table 1: Estimated covariate effects (posterior mean and 95% credible interval) in the administrative censoring group from the VCMs for the continuous CES-D data in the HERS.

|                     | VCM1                        | VCM2                        |
|---------------------|-----------------------------|-----------------------------|
| Intercept           | 20.63 [13.76,27.50]         | 20.47 [13.69,27.17]         |
| Black               | -5.51 [-9.93,-1.34]         | -5.37 [-9.98,-0.89]         |
| White               | -2.82 [-7.81,1.97]          | -2.68 [-7.58, 2.30]         |
| Baseline CD4        | 4.18 [-1.82,10.27]          | 4.22 [-1.94,10.46]          |
| Time                | -1.34 [-4.41,1.64]          | -1.35 [-4.23,1.54]          |
| Interaction between time and baseline CD4 | -0.53 [-3.65,2.64] | -0.58 [-3.63,2.49] |
Table 2: Estimated variance components (posterior mean and 95% credible interval) from the LMM and VCM1 for the continuous CES-D data in the HERS.

|                        | LMM          | VCM1         |
|------------------------|--------------|--------------|
| Random intercept SD    | 10.49 [9.85, 11.17] | 10.51 [9.85, 11.17] |
| Random slope SD        | 4.83 [4.32, 5.35]  | 4.85 [4.34, 5.37]  |
| Correlation between random effects | -0.17 [-0.22, -0.12] | -0.18 [-0.22, -0.13] |
| Error SD               | 8.19 [8.03, 8.35]  | 8.18 [8.03, 8.35]  |