A WinBUGS code for ignorable analysis in Section 5.1.1

model{
  for(i in 1:409){
    ## the conditional distribution of Y
    Y[i] ~ dnorm(mu[i], inv_tau)
    mu[i] <- alpha_0*(1-Z[i]) + alpha_1*Z[i] + beta_1[1]*X[i,1] + beta_1[2]*X[i,2] + beta_1[3]*X[i,3] + beta_1[4]*X[i,4]

    ## distribution of X including baseline score X[i,1] and indicators of center X[i,2:5]
    X[i,1] ~ dnorm(muX[i], inv_sigma)
    muX[i] <- alpha_00 + beta_0[1]*X[i,2] + beta_0[2]*X[i,3] + beta_0[3]*X[i,4]
    X[i,2:5] ~ dmulti(prX[1:4], 1)
  }

  ## conditional mean of Y for each Z
  avg_Y0 <- alpha_00 + beta_0[1]*prX[1] + beta_0[2]*prX[2] + beta_0[3]*prX[3]
  m0 <- alpha_0 + beta_1[1]*avg_Y0 + beta_1[2]*prX[1] + beta_1[3]*prX[2] + beta_1[4]*prX[3]
  m1 <- alpha_1 + beta_1[1]*avg_Y0 + beta_1[2]*prX[1] + beta_1[3]*prX[2] + beta_1[4]*prX[3]

  ## treatment effect
  theta <- alpha_1 - alpha_0

  ## priors
  for(k in 1:3){
    beta_0[k] ~ dnorm(0, 0.001)
    beta_1[k] ~ dnorm(0, 0.001)
  }
}
alpha_00 ~ dnorm(0, 0.001)
alpha_0 ~ dnorm(0, 0.001)
alpha_1 ~ dnorm(0, 0.001)
beta_1[4] ~ dnorm(0, 0.001)
inv_sigma ~ dgamma(0.01, 0.01)
inv_tau ~ dgamma(0.01, 0.01)
prX[1:4] ~ ddirich(a[])
}

B WinBUGS code for nonignorable analysis in Section 5.1.2

model{
for(i in 1:409){
  ## the conditional distribution of Y
  Y[i] ~ dnorm(mu[i], inv_tau)
  mu[i] <- alpha_0[R[i]]*(1-Z[i]) + alpha_1[R[i]]*Z[i] + beta_1[1]*X[i,1] + beta_1[2]*X[i,2]
  + beta_1[3]*X[i,3] + beta_1[4]*X[i,4]

  ## distribution of X including baseline score X[i,1] and indicators of center X[i,2:5]
  X[i,1] ~ dnorm(muX[i], inv_sigma)
  muX[i] <- alpha_00 + beta_0[1]*X[i,2] + beta_0[2]*X[i,3] + beta_0[3]*X[i,4]
  X[i,2:5] ~ dmulti(prX[1:4], 1)

  ## hazard model for the pattern
  for(j in 1:3){
    logit(h[i,j]) <- lambda_0[j] + lambda_Z*Z[i] + lambda_X[1]*X[i,1] + lambda_X[2]*X[i,2]
    + lambda_X[3]*X[i,3] + lambda_X[4]*X[i,4]
  }

  ## probabilities for being in each pattern
  R[i] ~ dcat(p[i,1:4])
  p[i,1] <- h[i,1]
  p[i,2] <- h[i,2]*(1 - p[i,1])
  p[i,3] <- h[i,3]*(1 - p[i,1] - p[i,2])
  p[i,4] <- 1 - sum(p[i,1:3])
}

## MC integration
## create L responses with covariates for each Z
for(i in 1:L){
  Z.mc[i] <- 0
}
\begin{verbatim}
Z_mc[i+L] <- 1

for(i in 1:L2){
  X_mc[i,1] ~ dnorm(muX_mc[i], inv_sigma)
  muX_mc[i] <- alpha_00 + beta_0[1]*X_mc[i,2] + beta_0[2]*X_mc[i,3] + beta_0[3]*X_mc[i,4]
  X_mc[i,2:5] ~ dmulti(prX[1:4], 1)

  ## conditional mean of Y for each pattern
  for(k in 1:4){
    mu_mc[i,k] <- alpha_0[k]*(1-Z_mc[i]) + alpha_1[k]*Z_mc[i] + beta_1[1]*X_mc[i,1]
    + beta_1[2]*X_mc[i,2] + beta_1[3]*X_mc[i,3] + beta_1[4]*X_mc[i,4]
  }

  ## hazard model
  for(j in 1:3){
    logit(h_mc[i, j]) <- lambda_0[j] + lambda_Z*Z_mc[i] + lambda_X[1]*X_mc[i,1]
    + lambda_X[2]*X_mc[i,2] + lambda_X[3]*X_mc[i,3] + lambda_X[4]*X_mc[i,4]
  }

  ## probabilities for being in each pattern
  p_mc[i,1] <- h_mc[i,1]
  p_mc[i,2] <- h_mc[i,2]*(1 - p_mc[i,1])
  p_mc[i,3] <- h_mc[i,3]*(1 - p_mc[i,1] - p_mc[i,2])
  p_mc[i,4] <- 1 - sum(p_mc[i,1:3])

  ## average over patterns
  avg_mu_mc[i] <- mu_mc[i,1]*p_mc[i,1] + mu_mc[i,2]*p_mc[i,2] + mu_mc[i,3]*p_mc[i,3]
    + mu_mc[i,4]*p_mc[i,4]
}

## treatment effect on means
m0 <- mean(avg_mu_mc[1:L])
ml <- mean(avg_mu_mc[(L+1):L2])
theta <- ml - m0

## priors for identified parameters
for(k in 1:3){
  alpha_0[k] ~ dnorm(0, 0.001)
  alpha_1[k] ~ dnorm(0, 0.001)
  beta_0[k] ~ dnorm(0, 0.001)
  lambda_0[k] ~ dnorm(0, 0.001)
}

for(j in 1:4){
  beta_1[j] ~ dnorm(0, 0.001)
  lambda_X[j] ~ dnorm(0, 0.001)
}
\end{verbatim}
\begin{verbatim}
alpha_00 ~ dnorm(0, 0.001)
lambda_Z ~ dnorm(0, 0.001)
inv_tau ~ dgamma(0.01, 0.01)
inv_sigma ~ dgamma(0.01, 0.01)
prX[1:4] ~ ddirich(a[])

## priors for unidentified parameters
alpha_0[4] <- zeta_00 + zeta_01*(K+C)
alpha_1[4] <- zeta_01 + zeta_11*(K+C)
zeta_00 <- inprod(M[1, ], alpha_0[1:3])
zeta_10 <- inprod(M[2, ], alpha_0[1:3])
zeta_01 <- inprod(M[1, ], alpha_1[1:3])
zeta_11 <- inprod(M[2, ], alpha_1[1:3])
\end{verbatim}

C WinBUGS code for two pattern model in Section 5.1.3

model{
for(i in 1:409){
## the conditional distribution of Y
Y[i] ~ dnorm(mu[i], inv_tau)
mu[i] <- alpha_0[R[i]]*(1-Z[i]) + alpha_1[R[i]]*Z[i] + beta_1[1]*X[i,1]
+ beta_1[2]*X[i,2] + beta_1[3]*X[i,3] + beta_1[4]*X[i,4]

## distribution of X including baseline score X[i,1] and indicators of center X[i,2:5]
X[i, 1] ~ dnorm(muX[i], inv_sigma)
muX[i] <- alpha_00 + beta_0[1]*X[i,2] + beta_0[2]*X[i,3] + beta_0[3]*X[i,4]
X[i, 2:5] ~ dmulti(prX[1:4], 1)

## hazard model
logit(h[i]) <- lambda_0 + lambda_Z*Z[i] + lambda_X[1]*X[i,1] + lambda_X[2]*X[i,2]
+ lambda_X[3]*X[i,3] + lambda_X[4]*X[i,4]

## probabilities for being in each pattern
R[i] ~ dcat(p[i, 1:2])
p[i, 1] <- h[i]
p[i, 2] <- 1 - h[i]
}

## MC integration
## create L responses with covariates for each Z
for(i in 1:L){
  Z_mc[i] <- 0
  Z_mc[i+L] <- 1
}

for(i in 1:L2){
  X_mc[i, 1] ~ dnorm(muX_mc[i], inv_sigma)
  muX_mc[i] <- alpha_00 + beta_0[1]*X_mc[i,2] + beta_0[2]*X_mc[i,3] + beta_0[3]*X_mc[i,4]
  X_mc[i, 2:5] ~ dmulti(prX[1:4], 1)
}

## conditional mean of Y for each pattern
for(k in 1:2){
  mu_mc[i, k] <- alpha_0[k]*(1-Z_mc[i]) + alpha_1[k]*Z_mc[i] + beta_1[1]*X_mc[i,1] + beta_1[2]*X_mc[i,2] + beta_1[3]*X_mc[i,3] + beta_1[4]*X_mc[i,4]
}

## hazard model
logit(h_mc[i]) <- lambda_0 + lambda_Z*Z_mc[i] + lambda_X[1]*X_mc[i,1] + lambda_X[2]*X_mc[i,2] + lambda_X[3]*X_mc[i,3] + lambda_X[4]*X_mc[i,4]

## probabilities for being in each pattern
p_mc[i, 1] <- h_mc[i]
p_mc[i, 2] <- 1 - h_mc[i]

## average over patterns
avg_mu_mc[i] <- mu_mc[i,1]*p_mc[i,1] + mu_mc[i,2]*p_mc[i,2]

## compute the mean of the L response Y2 for each Z
m0 <- mean(avg_mu_mc[1:L])
m1 <- mean(avg_mu_mc[(L+1):L2])
theta <- m1 - m0

## priors
alpha_0[1] ~ dnorm(0, 0.001)
alpha_1[1] ~ dnorm(0, 0.001)
alpha_0[2] <- alpha_0[1] + delta_0
alpha_1[2] <- alpha_1[1] + delta_1
for(k in 1:3){
beta_0[k] ~ dnorm(0, 0.001)
beta_1[k] ~ dnorm(0, 0.001)
lambda_X[k] ~ dnorm(0, 0.001)
}

lambda_0 ~ dnorm(0, 0.001)
alpha_00 ~ dnorm(0, 0.001)
beta_1[4] ~ dnorm(0, 0.001)
lambda_X[4] ~ dnorm(0, 0.001)
lambda_Z ~ dnorm(0, 0.001)
lambda_Y ~ dnorm(0, 0.001)
inv_sigma ~ dgamma(0.01, 0.01)
inv_tau ~ dgamma(0.01, 0.01)
prX[1:4] ~ ddirich(a[])

D Estimates of models in Sections 5.1.2 and 5.2
Table 1: Posterior summaries for observed data parameters in the RAM-PMM in Section 5.1.2.

| parameter | mean  | 95% CI     |
|-----------|-------|------------|
| $\alpha_0^{(1)}$ | 23.7  | (18.7, 28.5) |
| $\alpha_0^{(2)}$ | 23.6  | (18.2, 29.1) |
| $\alpha_0^{(3)}$ | 19.8  | (12.9, 26.7) |
| $\alpha_1^{(1)}$ | 23.8  | (19.03, 28.5) |
| $\alpha_1^{(2)}$ | 22.5  | (16.8, 28.1) |
| $\alpha_1^{(3)}$ | 20.4  | (12.9, 28.1) |
| $\beta_1[1]$ | 0.5   | (0.4, 0.6)  |
| $\beta_1[2]$ | -0.8  | (-5.1, 3.5) |
| $\beta_1[3]$ | -4.1  | (-7.3, -0.9) |
| $\beta_1[4]$ | 0.3   | (-3.8, 4.4) |
| $\lambda_0$ | 0.1   | (-0.7, 0.8) |
| $\lambda_1$ | 2.1   | (1.3, 2.98) |
| $\lambda_2$ | 0.7   | (-0.2, 1.6) |
| $\lambda_z$ | -0.3  | (-0.6, 0.01) |
| $\lambda_x[1]$ | 0.002 | (-0.01, 0.02) |
| $\lambda_x[2]$ | -1.5  | (-2.04, -1.02) |
| $\lambda_x[3]$ | 0.5   | (-0.1, 0.99) |
| $\lambda_x[4]$ | -1.3  | (-1.7, -0.8) |

Table 2: Posterior summaries for parameters which are a function of the sensitivity parameter $C$ in the RAM-PMM in Section 5.1.2.

| $C$ | 0     | 1     | 2     | 3     |
|-----|-------|-------|-------|-------|
| parameter | mean  | 95% CI | mean  | 95% CI | mean  | 95% CI | mean  | 95% CI |
| $\alpha_0^{(4)}$ | 20.5  | (13.8, 27.1) | 18.6  | (9.5, 27.5) | 16.7  | (5.1, 28.2) | 14.98 | (0.8, 29.3) |
| $\alpha_1^{(4)}$ | 20.5  | (13.4, 27.7) | 18.9  | (8.8, 28.8) | 17.2  | (4.1, 30.3) | 15.7  | (-0.7, 32.2) |
Table 3: Parameter estimates in the RAM-SM response (linear regression) model.

| parameter | estimate | 95% CI       |
|-----------|----------|--------------|
| $\beta_0^*$ | 23.3     | (18.3, 28.3) |
| $\theta$    | -1.5     | (-3.8, 0.8)  |
| $\beta^*[1]$ | 0.5      | (0.4, 0.6)   |
| $\beta^*[2]$ | -1.1     | (-4.5, 2.2)  |
| $\beta^*[3]$ | -3.2     | (-6.4, 0.1)  |
| $\beta^*[4]$ | 0.9      | (-2.3, 4.1)  |

Table 4: Parameter estimates in the RAM-SM missing data mechanism.

| parameter | estimate | 95% CI       |
|-----------|----------|--------------|
| $\lambda_{01}$ | -0.6     | (-1.8, 0.6)  |
| $\lambda_{02}$ | 1.7      | (0.5, 2.9)   |
| $\lambda_{03}$ | 0.5      | (-0.7, 1.7)  |
| $\gamma$    | -1.2     | (-2.6, 0.2)  |
| $\lambda[1]$ | -0.02    | (-0.04, -0.003) |
| $\lambda[2]$ | -1.7     | (-2.3, -1.1) |
| $\lambda[3]$ | 0.7      | (0.1, 1.2)   |
| $\lambda[4]$ | -1.5     | (-2.1, -0.9) |
| $\delta_1$  | 0.04     | (0.01, 0.1)  |
| $\delta_2$  | 0.03     | (-0.01, 0.1) |