

Bayesian Joint Models, with Application to HIV Data (R Code)

Harper Xiaolan Cheng

1. Data Processing

```
hiv <- read_csv("~/Desktop/UBC/project/rebound_data.csv")

dat0 <- hiv %>%
  drop_na(RNA_V, CD4_V) %>%
  group_by(PATIENT) %>%
  mutate(first=as.numeric(row_number()==min(row_number()[RNA_V==1]))) %>%
  filter(first==1)

dat1 <- hiv %>%
  drop_na(RNA_V, CD4_V) %>%
  filter(RNA_V!=-1) %>%
  full_join(dat0) %>%
  arrange(PATIENT, days_from_seroco)

dat <- dat1 %>%
  mutate(RNA_V=if_else(RNA_V==1, log10(RNA_L/2), log10(RNA_V)),
         CD4_V=log(CD4_V),
         months_from_seroco=days_from_seroco/30,
         art_start=art_start/30,
         ati_start=ati_start/30) %>%
  dplyr::select(-RNA_T, -RNA_L, -days_from_seroco, -first) %>%
  filter(is.na(months_from_seroco)==FALSE)
```

2. Define Functions

```
# 2.1 Find RNA_V viral rebound
Find_rebound <- function(rna) {
  # Find the index of viral rebound. If there is no rebound yet, then return 0.
  if (length(rna)==1) ind <- 0
  else {
    for (i in 2:length(rna)) {
      if (rna[i] > rna[i-1]) {ind <- i; break}
      else {ind <- 0}
    }
  }
  return(ind)
}
```

```

# 2.2 Find RNA_V peak points
Find_peak <- function(rna) {
  # Find the index of viral peak point. If there is no peak yet, then return 0.
  ind <- which.max(rna)
  if (ind == length(rna)) ind <- 0
  return(ind)
}

# 2.3 Define `status` of viral rebound/viral peak
Define_status <- function(fn, rna) {
  # Define the status: 1 - there is a rebound or peak; 0 - there is no rebound or peak
  # Returns `status` and `ind`
  ind <- fn(rna)
  if (ind!=0) status <- 1
  else {status <- 0; ind <- length(rna)}
  df <- c(status, ind)
  return(df)
}

```

3. Select Eligible Patients

```

# 3.1 LME data
lme.dat <- dat %>%
  mutate(t1=months_from_seroco-art_start) %>%
  filter(treatment==1 & t1<6 & t1!=0) %>%
  group_by(PATIENT) %>%
  filter(length(t1)>1)
pid1 <- unique(lme.dat$PATIENT)

# 3.2 survival data
## 3.2.1 RNA rebound
surv.dat.rebound <- dat %>%
  filter(months_from_seroco >= ati_start) %>%
  group_by(PATIENT) %>%
  mutate(status=Define_status(Find_rebound, RNA_V)[1],
         t2=months_from_seroco-art_start,
         t3=months_from_seroco-ati_start) %>%
  slice(Define_status(Find_rebound, RNA_V)[2])
pid2 <- surv.dat.rebound$PATIENT

## 3.2.2 RNA peak point:
surv.dat.peak <- dat %>%
  filter(months_from_seroco >= ati_start) %>%
  group_by(PATIENT) %>%
  mutate(status=Define_status(Find_peak, RNA_V)[1],
         t2=months_from_seroco-art_start,
         t3=months_from_seroco-ati_start) %>%
  slice(Define_status(Find_peak, RNA_V)[2])
pid3 <- surv.dat.peak$PATIENT

```

```

# 3.3 find common patients
pid <- Reduce(intersect, list(pid1, pid2, pid3))
length(pid)

lme.dat <- lme.dat %>% filter(PATIENT %in% pid)
surv.dat.peak <- surv.dat.peak %>% filter(PATIENT %in% pid)
surv.dat.rebound <- surv.dat.rebound %>% filter(PATIENT %in% pid)

# 3.4 identify outliers
lme.dat %>% filter(PATIENT==23) %>% summarise(m=10^mean(RNA_V))
lme.dat %>% filter(PATIENT!=23) %>% ungroup() %>% summarise(m=10^mean(RNA_V))
lme.dat %>% filter(PATIENT==50) %>% summarise(m=exp(mean(CD4_V)))
lme.dat %>% filter(PATIENT!=50) %>% ungroup() %>% summarise(m=exp(mean(CD4_V)))

# 3.5 remove outliers
lme.dat.rm <- lme.dat %>% filter(!(PATIENT %in% c(23, 50)))
surv.dat.peak.rm <- surv.dat.peak %>% filter(!(PATIENT %in% c(23, 50)))
surv.dat.rebound.rm <- surv.dat.rebound %>% filter(!(PATIENT %in% c(23, 50)))

```

4. Exploratory Analysis

```

# 4.1 RNA viral loads trajectories
dat %>%
  ggplot(aes(x=months_from_seroco-art_start, y=RNA_V, group=PATIENT)) +
  geom_line(alpha=0.5, color="dark blue") +
  geom_point(alpha=0.2, size=0.8) +
  xlim(0,100) +
  labs(title="Individual Trajectories of (log10) RNA Viral Loads on Eligible Patients",
        subtitle="For the Duration of the Study",
        x="Months from the Start of ART", y="log10 of RNA Viral Load") +
  theme_bw()
ggsave("EDA_RNA.pdf", width=6, height=5)

# 4.2 CD4 cell counts trajectories
dat %>%
  ggplot(aes(x=months_from_seroco-art_start, y=CD4_V, group=PATIENT)) +
  geom_line(alpha=0.5, color="dark blue") +
  geom_point(alpha=0.2, size=0.8) +
  xlim(0,100) +
  labs(title="Individual Trajectories of (log) CD4 Cell Counts on Eligible Patients",
        subtitle="For the Duration of the Study",
        x="Months from the Start of ART", y="Natural log of CD4 Counts") +
  theme_bw()
ggsave("EDA_CD4.pdf", width=6, height=5)

```

5. Survival analysis (Cox Model)

```
# 5.1 Model Selection
## 5.1.1 RNA rebound
rebound.cox1 <- coxph(Surv(t2, status)~GENDER+age,
                      data=surv.dat.rebound.rm, x=TRUE)
rebound.cox2 <- coxph(Surv(t2, status)~1,
                      data=surv.dat.rebound.rm, x=TRUE)
anova(rebound.cox1, rebound.cox2)

## 5.1.2 RNA peak point
peak.cox1 <- coxph(Surv(t2, status)~GENDER+age,
                   data=surv.dat.peak.rm, x=TRUE)
peak.cox2 <- coxph(Surv(t2, status)~1,
                   data=surv.dat.peak.rm, x=TRUE)
anova(peak.cox1, peak.cox2)

# 5.2 COX Model on rebound and peak point
cox.fit.rebound <- coxph(Surv(t2, status)~1, data=surv.dat.rebound.rm, x=TRUE)
cox.fit.peak <- coxph(Surv(t2, status)~1, data=surv.dat.peak.rm, x=TRUE)

# 5.3 Histogram
surv.dat.rebound.rm %>%
  ggplot(aes(x=t3)) +
  geom_histogram(color="black", fill="gray", bins=60) +
  theme_bw() +
  xlim(0,30) +
  labs(title="Histogram on Time to Viral Load Rebound",
       x="Months since ART Interruption", y="Number of Patients")
ggsave("hist_rebound.pdf", width=6, height=5)

surv.dat.peak.rm %>%
  ggplot(aes(x=t3)) +
  geom_histogram(color="black", fill="gray", bins=60) +
  theme_bw() +
  xlim(0,100) +
  labs(title="Histogram on Time to Viral Load Peak Point",
       x="Months since ART Interruption", y="Number of Patients")
ggsave("hist_peak.pdf", width=6, height=5)
```

6. LME

```
# 6.1 Model Selection
## 6.1.1 RNA viral load decay
rna.lme1 <- lme(RNA_V~t1+age+GENDER, random=~t1|PATIENT, data=lme.dat, method="ML")
rna.lme2 <- lme(RNA_V~t1, random=~t1|PATIENT, data=lme.dat, method="ML")
anova(rna.lme1, rna.lme2)

## 6.1.2 CD4 cell counts
```

```

cd4.lme1 <- lme(CD4_V~t1+age+GENDER, random=~t1|PATIENT, data=lme.dat, method="ML")
cd4.lme2 <- lme(CD4_V~t1, random=~t1|PATIENT, data=lme.dat, method="ML")
anova(cd4.lme1, cd4.lme2)

# 6.2 LME Model
## 6.2.1 RNA viral loads
lme.fit.rna <- lme(RNA_V~t1, random=~t1|PATIENT, data=lme.dat)
lme.fit.rna.rm <- lme(RNA_V~t1, random=~t1|PATIENT, data=lme.dat.rm)

## 6.2.2 CD4 cell counts
lme.fit.cd4 <- lme(CD4_V~t1, random=~t1|PATIENT, data=lme.dat)
lme.fit.cd4.rm <- lme(CD4_V~t1, random=~t1|PATIENT, data=lme.dat.rm)

# 6.3 Individual Trajectories
## 6.3.1 RNA viral loads
lme.dat.rm %>%
  ggplot(aes(x=t1, y=RNA_V, group=PATIENT)) +
    geom_line(alpha=0.5, color="dark blue") +
    geom_point(alpha=0.2, size=0.8) +
    ylim(0,6) +
    labs(title="Individual Trajectories of (log10) RNA Viral Loads on Eligible Patients",
         subtitle="First 6 months of ART Treatment",
         x="Months from the Start of ART", y="RNA Viral Load (log10 scale)") +
    theme_bw()
ggsave("long_RNA.pdf", width=6, height=5)

## 6.3.2 CD4 cell counts
lme.dat.rm %>%
  ggplot(aes(x=t1, y=CD4_V, group=PATIENT)) +
    geom_line(alpha=0.5, color="dark blue") +
    geom_point(alpha=0.2, size=0.8) +
    ylim(5,7.5) +
    labs(title="Individual Trajectories of (log) CD4 Cell Counts on Eligible Patients",
         subtitle="First 6 months of ART Treatment",
         x="Months from the Start of ART", y="Natural log of CD4 Counts") +
    theme_bw()
ggsave("long_CD4.pdf", width=6, height=5)

# 6.4 Diagnostics
## 6.4.1 residual plot
pdf("lme_residual1.pdf", width=6, height=5)
plot(lme.fit.rna.rm, resid(.,type="p")~fitted(.,id=0.05,adj=-0.3)
dev.off()
pdf("lme_residual2.pdf", width=6, height=5)
plot(lme.fit.cd4.rm, resid(.,type="p")~fitted(.,id=0.05,adj=-0.3)
dev.off()

## 6.4.2 QQ plot for error terms
pdf("lme_error1.pdf", width=6, height=5)
qqnorm(lme.fit.rna.rm, ~resid(.))

```

```

dev.off()
pdf("lme_error2.pdf", width=6, height=5)
qqnorm(lme.fit.cd4.rm, ~resid(.))
dev.off()

## 6.4.3 QQ plot for random effects
pdf("lme_ranef1.pdf", width=6, height=5)
qqnorm(lme.fit.rna.rm, ~ranef(.), id=0.1, cex=0.7)
dev.off()
pdf("lme_ranef2.pdf", width=6, height=5)
qqnorm(lme.fit.cd4.rm, ~ranef(.), id=0.1, cex=0.7)
dev.off()

# 6.5 Goodness of fit
## 6.5.1 RNA viral loads
lme.dat.rm %>%
  add_column(pred1=predict(lme.fit.rna.rm), pred2=predict(lme.fit.cd4.rm)) %>%
  mutate(patid=as.factor(PATIENT)) %>%
  filter(patid %in% c(1, 29, 36, 43, 69, 75)) %>%
  ggplot(aes(x=t1, y=pred1, group=patid)) +
  geom_line() +
  geom_point(aes(x=t1, y=RNA_V, group=patid), size=2, alpha=0.5) +
  facet_wrap(~patid) +
  theme_bw() +
  labs(title="LME Prediction Plot for (log10) RNA Viral Loads on Selected Patients",
        x="Time in Month", y="RNA Viral Loads on log10 Scale")
ggsave("rna_pred.pdf", width=7, height=5)

## 6.5.2 CD4 cell counts
lme.dat.rm %>%
  add_column(pred1=predict(lme.fit.rna.rm), pred2=predict(lme.fit.cd4.rm)) %>%
  mutate(patid=as.factor(PATIENT)) %>%
  filter(patid %in% c(1, 29, 36, 43, 69, 75)) %>%
  ggplot(aes(x=t1, y=pred2, group=patid)) +
  geom_line() +
  geom_point(aes(x=t1, y=CD4_V, group=patid), size=2, alpha=0.5) +
  facet_wrap(~patid) +
  theme_bw() +
  labs(title="LME Prediction Plot for (log) CD4 Cell Counts on Selected Patients",
        x="Time in Month", y="CD4 Cell Counts on Natural Log Scale")
ggsave("cd4_pred.pdf", width=7, height=5)

```

7. Bayesian Joint model (shared random effects)

```

# 7.1 Model analysis
## 7.1.1 RNA viral load and time to rebound
joint.fit1 <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                             param="shared-RE", n.iter=50000)
summary(joint.fit1)

```

```

## 7.1.2 RNA viral load and time to peak points
joint.fit2 <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                             param="shared-RE", n.iter=50000)
summary(joint.fit2)

## 7.1.3 CD4 cell counts and time to rebound
joint.fit3 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                             param="shared-RE", n.iter=50000)
summary(joint.fit3)

## 7.1.4 CD4 cell counts and time to peak points
joint.fit4 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                             param="shared-RE", n.iter=50000)
summary(joint.fit4)

```

7.2 Diagnostics for Analysis 1

```

## 7.2.1 Trace plot
pp <- do.call(cbind, joint.fit1$mcmc["alphas"]) %>% data.frame()

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,1]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Intercept Association Structure  $\alpha^{(r)}_{a_0}$ "),
        subtitle="Analysis 1: Joint Model on Viral Load and Viral Rebound
                  with Shared Random Effects",
        x="iteration", y=TeX(" $\alpha^{(r)}_{a_0}$ ")) +
  theme_bw()
ggsave("trace1.pdf", width=7, height=3.5)

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,2]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Slope Association Structure  $\alpha^{(r)}_{a_1}$ "),
        subtitle="Analysis 1: Joint Model on Viral Load and Viral Rebound
                  with Shared Random Effects",
        x="iteration", y=TeX(" $\alpha^{(r)}_{a_1}$ ")) +
  theme_bw()
ggsave("trace2.pdf", width=7, height=3.5)

## 7.2.2 ACF plot
pp1 <- with(acf(pp[,1], plot=FALSE), data.frame(lag, acf))
pp2 <- with(acf(pp[,2], plot=FALSE), data.frame(lag, acf))

ggplot(aes(x=lag, y=acf), data=pp1) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Intercept Association Structure
                   $\alpha^{(r)}_{a_0}$ "),
        subtitle="Analysis 1: Joint Model on Viral Load and Viral Rebound
                  with Shared Random Effects",
        x="lag", y="Autocorrelation") +
  theme_bw()

```

```

ggsave("acf1.pdf", width=7, height=3.5)

ggplot(aes(x=lag, y=acf), data=pp2) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Slope Association Structure
     $\alpha^{(r)}_{a_1}$ "),
    subtitle="Analysis 1: Joint Model on Viral Load and Viral Rebound
    with Shared Random Effects",
    x="lag", y="Autocorrelation") +
  theme_bw()
ggsave("acf2.pdf", width=7, height=3.5)

```

7.3 Diagnostics for Analysis 2-4

```

### 7.3.2 Analysis 2
### Trace plot
pp <- do.call(cbind, joint.fit2$mcmc["alphas"])) %>% data.frame()

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,1]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Intercept Association Structure
     $\alpha^{(p)}_{a_0}$ "),
    subtitle="Analysis 2: Joint Model on Viral Loads and Viral Peak Point
    with Shared Random Effects",
    x="iteration", y=TeX(" $\alpha^{(p)}_{a_0}$ ")) +
  theme_bw()
ggsave("trace4.pdf", width=7, height=3.5)

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,2]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Slope Association Structure  $\alpha^{(p)}_{a_1}$ "),
    subtitle="Analysis 2: Joint Model on Viral Loads and Viral Peak Point
    with Shared Random Effects",
    x="iteration", y=TeX(" $\alpha^{(p)}_{a_1}$ ")) +
  theme_bw()
ggsave("trace5.pdf", width=7, height=3.5)

### ACF plot
pp <- do.call(cbind, joint.fit2$mcmc["alphas"])) %>% data.frame()
pp1 <- with(acf(pp[,1], plot=FALSE), data.frame(lag, acf))
pp2 <- with(acf(pp[,2], plot=FALSE), data.frame(lag, acf))

ggplot(aes(x=lag, y=acf), data=pp1) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Intercept Association Structure
     $\alpha^{(p)}_{a_0}$ "),
    subtitle="Analysis 2: Joint Model on Viral Loads and Viral Peak Point
    with Shared Random Effects",
    x="lag", y="Autocorrelation") +
  theme_bw()

```



```

ggsave("acf4.pdf", width=7, height=3.5)

ggplot(aes(x=lag, y=acf), data=pp2) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Slope Association Structure
     $\alpha^{(p)}_{a_1}$ "),
    subtitle="Analysis 2: Joint Model on Viral Loads and Viral Peak Point
    with Shared Random Effects",
    x="lag", y="Autocorrelation") +
  theme_bw()
ggsave("acf5.pdf", width=7, height=3.5)

### 7.3.3 Analysis 3
### Trace plot
pp <- do.call(cbind, joint.fit3$mcmc["alphas"])) %>% data.frame()

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,1]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Intercept Association Structure  $\alpha^{(r)}_{b_0}$ "),
    subtitle="Analysis 3: Joint Model on CD4 Cell Counts and Viral Rebound
    with Shared Random Effects",
    x="iteration", y=TeX(" $\alpha^{(r)}_{b_0}$ ")) +
  theme_bw()
ggsave("trace6.pdf", width=7, height=3.5)

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,2]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Slope Association Structure  $\alpha^{(r)}_{b_1}$ "),
    subtitle="Analysis 3: Joint Model on CD4 Cell Counts and Viral Rebound
    with Shared Random Effects",
    x="iteration", y=TeX(" $\alpha^{(r)}_{b_1}$ ")) +
  theme_bw()
ggsave("trace7.pdf", width=7, height=3.5)

### ACF plot
pp <- do.call(cbind, joint.fit3$mcmc["alphas"])) %>% data.frame()
pp1 <- with(acf(pp[,1], plot=FALSE), data.frame(lag, acf))
pp2 <- with(acf(pp[,2], plot=FALSE), data.frame(lag, acf))

ggplot(aes(x=lag, y=acf), data=pp1) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Intercept Association Structure
     $\alpha^{(r)}_{b_0}$ "),
    subtitle="Analysis 3: Joint Model on CD4 Cell Counts and Viral Rebound
    with Shared Random Effects",
    x="lag", y="Autocorrelation") +
  theme_bw()
ggsave("acf6.pdf", width=7, height=3.5)

ggplot(aes(x=lag, y=acf), data=pp2) +
  geom_hline(aes(yintercept=0)) +

```

```

geom_segment(aes(xend=lag, yend=0)) +
labs(title=TeX("Autocorrelation Plot for Slope Association Structure  $\alpha_{b_1}^{(r)}$ "),
      subtitle="Analysis 3: Joint Model on CD4 Cell Counts and Viral Rebound
               with Shared Random Effects",
      x="lag", y="Autocorrelation") +
theme_bw()
ggsave("acf7.pdf", width=7, height=3.5)

### 7.3.4 Analysis 4
### Trace plot
pp <- do.call(cbind, joint.fit4$mcmc["alphas"])) %>% data.frame()

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,1]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Intercept Association Structure  $\alpha_{b_0}^{(p)}$ "),
        subtitle="Analysis 4: Joint Model on CD4 Cell Counts and Viral Peak Point
                  with Shared Random Effects",
        x="iteration", y=TeX(" $\alpha_{b_0}^{(p)}$ ")) +
  theme_bw()
ggsave("trace8.pdf", width=7, height=3.5)

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,2]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Slope Association Structure  $\alpha_{b_1}^{(p)}$ "),
        subtitle="Analysis 4: Joint Model on CD4 Cell Counts and Viral Peak Point
                  with Shared Random Effects",
        x="iteration", y=TeX(" $\alpha_{b_1}^{(p)}$ ")) +
  theme_bw()
ggsave("trace9.pdf", width=7, height=3.5)

### ACF plot
pp <- do.call(cbind, joint.fit4$mcmc["alphas"])) %>% data.frame()
pp1 <- with(acf(pp[,1], plot=FALSE), data.frame(lag, acf))
pp2 <- with(acf(pp[,2], plot=FALSE), data.frame(lag, acf))

ggplot(aes(x=lag, y=acf), data=pp1) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Intercept Association Structure
                   $\alpha_{b_0}^{(p)}$ "),
        subtitle="Analysis 4: Joint Model on CD4 Cell Counts and Viral Peak Point
                  with Shared Random Effects",
        x="lag", y="Autocorrelation") +
  theme_bw()
ggsave("acf8.pdf", width=7, height=3.5)

ggplot(aes(x=lag, y=acf), data=pp2) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Slope Association Structure
                   $\alpha_{b_1}^{(p)}$ "),
        subtitle="Analysis 4: Joint Model on CD4 Cell Counts and Viral Peak Point
                  with Shared Random Effects",

```

```

      x="lag", y="Autocorrelation") +
    theme_bw()
ggsave("acf9.pdf", width=7, height=3.5)

```

7.4 Sensitivity Analysis 1

```

## 7.4.1 RNA viral loads and time to rebound
joint.fit1.sen <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                                param="shared-RE", n.iter=50000,
                                control=list(priorVar=10))

summary(joint.fit1.sen)

## 7.4.2 RNA viral loads and time to peak point
joint.fit2.sen <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                                param="shared-RE", n.iter=50000,
                                control=list(priorVar=10))

summary(joint.fit2.sen)

## 7.4.3 CD4 cell counts and time to rebound
joint.fit3.sen <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                                param="shared-RE", n.iter=50000,
                                control=list(priorVar=10))

summary(joint.fit3.sen)

## 7.4.4 CD4 cell counts and time to peak point
joint.fit4.sen <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                                param="shared-RE", n.iter=50000,
                                control=list(priorVar=10))

summary(joint.fit4.sen)

```

7.5. Sensitivity Analysis 2

```

## 7.5.1 RNA viral loads and time to rebound
joint.fit1.sen2 <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                                param="shared-RE", n.iter=50000,
                                control=list(priorVar=1000))

summary(joint.fit1.sen2)

## 7.5.2 RNA viral loads and time to peak point
joint.fit2.sen2 <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                                param="shared-RE", n.iter=50000,
                                control=list(priorVar=1000))

summary(joint.fit2.sen2)

## 7.5.3 CD4 cell counts and time to rebound
joint.fit3.sen2 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                                param="shared-RE", n.iter=50000,
                                control=list(priorVar=1000))

summary(joint.fit3.sen2)

```

```
## 7.5.4 CD4 cell counts and time to peak point
joint.fit4.sen2 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                                param="shared-RE", n.iter=50000,
                                control=list(priorVar=1000))

summary(joint.fit4.sen2)
```

8. Bayesian Joint model (shared $\eta_i(t)$)

```
# 8.1 Model fitting
## 8.1.1 RNA viral load and time to rebound
joint.fit1.eta <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound,
                                timeVar="t1", n.iter=50000)

summary(joint.fit1.eta)

## 8.1.2 RNA viral load and time to peak points
joint.fit2.eta <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak,
                                timeVar="t1", n.iter=50000)

summary(joint.fit2.eta)

## 8.1.3 CD4 cell counts and time to rebound
joint.fit3.eta <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound,
                                timeVar="t1", n.iter=50000)

summary(joint.fit3.eta)

## 8.1.4 CD4 counts and time to peak points
joint.fit4.eta <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak,
                                timeVar="t1", n.iter=50000)

summary(joint.fit4.eta)
```

8.2 Sensitivity Analysis 3

```
# 8.2.1 RNA viral loads and time to rebound
joint.fit1.eta.sen <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                                    n.iter=50000, control=list(priorVar=1000))

summary(joint.fit1.eta.sen)

# 8.2.2 RNA viral loads and time to peak point
joint.fit2.eta.sen <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                                    n.iter=50000, control=list(priorVar=1000))

summary(joint.fit2.eta.sen)

# 8.2.3 CD4 cell counts and time to rebound
joint.fit3.eta.sen <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                                    n.iter=50000, control=list(priorVar=1000))

summary(joint.fit3.eta.sen)

# 8.2.4 CD4 cell counts and time to peak point
joint.fit4.eta.sen <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                                    n.iter=50000, control=list(priorVar=1000))

summary(joint.fit4.eta.sen)$`CoefTable-Event`[-3, -6] %>% xtable(digits=4)
```

8.3 Sensitivity Analysis 4

```
# 8.3.1 RNA viral loads and time to rebound
joint.fit1.eta.sen2 <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                                     n.iter=50000, control=list(priorVar=10))
summary(joint.fit1.eta.sen2)

# 8.3.2 RNA viral loads and time to peak point
joint.fit2.eta.sen2 <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                                     n.iter=50000, control=list(priorVar=10))
summary(joint.fit2.eta.sen2)

# 8.3.3 CD4 cell counts and time to rebound
joint.fit3.eta.sen2 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                                     n.iter=50000, control=list(priorVar=10))
summary(joint.fit3.eta.sen2)

# 8.3.4 CD4 cell counts and time to peak point
joint.fit4.eta.sen2 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                                     n.iter=50000, control=list(priorVar=10))
summary(joint.fit4.eta.sen2)
```

9. Bayesian Joint Model (shared $\eta_i(t)$ and $\eta'_i(t)$)

```
# 9.1 Model fitting
## 9.1.1 RNA viral loads and time to rebound
dForm <- list(fixed=~1, random=~1, indFixed=2, indRandom=2)
joint.fit1.eta.p <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                                   n.iter=50000, param="td-both", extraForm=dForm)
summary(joint.fit1.eta.p)

## 9.1.2 RNA viral loads and time to peak point
joint.fit2.eta.p <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                                   n.iter=50000, param="td-both", extraForm=dForm)
summary(joint.fit2.eta.p)

## 9.1.3 CD4 cell counts and time to rebound
joint.fit3.eta.p <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                                   n.iter=50000, param="td-both", extraForm=dForm)
summary(joint.fit3.eta.p)

## 9.1.4 CD4 cell counts and time to peak point
joint.fit4.eta.p <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                                   n.iter=50000, param="td-both", extraForm=dForm)
summary(joint.fit4.eta.p)
```

9.2 Sensitivity analysis 5

```
## 9.2.1 RNA viral loads and time to rebound
joint.fit1.eta.p.sen <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                                       n.iter=50000, param="td-both", extraForm=dForm,
                                       control=list(priorVar=1000))

summary(joint.fit1.eta.p.sen)

## 9.2.2 RNA viral loads and time to peak point
joint.fit2.eta.p.sen <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                                       n.iter=50000, param="td-both", extraForm=dForm,
                                       control=list(priorVar=1000))

summary(joint.fit2.eta.p.sen)

## 9.2.3 CD4 cell counts and time to rebound
joint.fit3.eta.p.sen <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                                       n.iter=50000, param="td-both", extraForm=dForm,
                                       control=list(priorVar=1000))

summary(joint.fit3.eta.p.sen)

## 9.2.4 CD4 cell counts and time to peak point
joint.fit4.eta.p.sen <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                                       n.iter=50000, param="td-both", extraForm=dForm,
                                       control=list(priorVar=1000))

summary(joint.fit4.eta.p.sen)
```

9.3 Sensitivity Analysis 6

```
## 9.3.1 RNA viral loads and time to rebound
joint.fit1.eta.p.sen2 <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                                       n.iter=50000, param="td-both", extraForm=dForm,
                                       control=list(priorVar=10))

summary(joint.fit1.eta.p.sen2)

## 9.3.2 RNA viral loads and time to peak point
joint.fit2.eta.p.sen2 <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                                       n.iter=50000, param="td-both", extraForm=dForm,
                                       control=list(priorVar=10))

summary(joint.fit2.eta.p.sen2)

## 9.3.3 CD4 cell counts and time to rebound
joint.fit3.eta.p.sen2 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                                       n.iter=50000, param="td-both", extraForm=dForm,
                                       control=list(priorVar=10))

summary(joint.fit3.eta.p.sen2)

## 9.3.4 CD4 cell counts and time to peak point
joint.fit4.eta.p.sen2 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                                       n.iter=50000, param="td-both", extraForm=dForm,
                                       control=list(priorVar=10))

summary(joint.fit4.eta.p.sen2)
```

10. Bayesian Joint Model (shared cumulative effects)

```
# 10.1 Model fitting
## 10.1.1 RNA viral loads and time to rebound
iForm <- list(fixed=~0+t1+I(t1^2/2), random=~0+t1+I(t1^2/2), indFixed=1:2, indRandom=1:2)
joint.fit1.cum <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                                n.iter=50000, param="td-extra", extraForm=iForm)
summary(joint.fit1.cum)

## 10.1.2 RNA viral loads and time to peak point
joint.fit2.cum <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                                n.iter=50000, param="td-extra", extraForm=iForm)
summary(joint.fit2.cum)

## 10.1.3 CD4 cell counts and time to rebound
joint.fit3.cum <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                                n.iter=50000, param="td-extra", extraForm=iForm)
summary(joint.fit3.cum)

## 10.1.4 CD4 cell counts and time to peak point
joint.fit4.cum <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                                n.iter=50000, param="td-extra", extraForm=iForm)
summary(joint.fit4.cum)
```

10.2 Diagnostics for Analysis 13

```
### 10.2.1 Trace plot
pp <- do.call(cbind, joint.fit1.cum$mcmc["Dalphas"])) %>% data.frame()

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,1]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Cumulative Association Structure  $\alpha^{(r)}_{y}$ "),
       subtitle="Analysis 13: Joint Model on Viral Load and Viral Rebound
               with Shared Cumulative Effects",
       x="iteration", y=TeX(" $\alpha^{(r)}_{y}$ ")) +
  theme_bw()
ggsave("trace3.pdf", width=7, height=3.5)

### 10.2.2 ACF plot
pp1 <- with(acf(pp[,1], plot=FALSE), data.frame(lag, acf))

ggplot(aes(x=lag, y=acf), data=pp1) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Cumulative Association Structure
                $\alpha^{(r)}_{y}$ "),
       subtitle="Analysis 13: Joint Model on Viral Load and Viral Rebound
               with Shared Cumulative Effects",
       x="lag", y="Autocorrelation") +
  theme_bw()
ggsave("acf3.pdf", width=7, height=4)
```

10.3 Diagnostics for Analysis 14-16

```
## 10.3.1 Analysis 14
### Trace plot
pp <- do.call(cbind, joint.fit2.cum$mcmc["Dalphas"]) %>% data.frame()

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,1]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Cumulative Association Structure  $\alpha^{(p)}_y$ "),
        subtitle="Analysis 14: Joint Model on Viral Loads and Viral Peak Point
                  with Shared Cumulative Effects",
        x="iteration", y=TeX(" $\alpha^{(p)}_y$ ")) +
  theme_bw()
ggsave("trace10.pdf", width=7, height=3.5)

### ACF plot
pp1 <- with(acf(pp[,1], plot=FALSE), data.frame(lag, acf))

ggplot(aes(x=lag, y=acf), data=pp1) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Cumulative Association Structure
                   $\alpha^{(p)}_y$ "),
        subtitle="Analysis 14: Joint Model on Viral Loads and Viral Peak Point
                  with Shared Cumulative Effects",
        x="lag", y="Autocorrelation") +
  theme_bw()
ggsave("acf10.pdf", width=7, height=4)

## 10.3.2 Analysis 15
### Trace plot
pp <- do.call(cbind, joint.fit3.cum$mcmc["Dalphas"]) %>% data.frame()

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,1]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Cumulative Association Structure  $\alpha^{(r)}_z$ "),
        subtitle="Analysis 15: Joint Model on CD4 Cell Counts and Viral Rebound
                  with Shared Cumulative Effects",
        x="iteration", y=TeX(" $\alpha^{(r)}_z$ ")) +
  theme_bw()
ggsave("trace11.pdf", width=7, height=3.5)

### ACF plot
pp1 <- with(acf(pp[,1], plot=FALSE), data.frame(lag, acf))

ggplot(aes(x=lag, y=acf), data=pp1) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Cumulative Association Structure
                   $\alpha^{(r)}_z$ "),
        subtitle="Analysis 15: Joint Model on CD4 Cell Counts and Viral Rebound
                  with Shared Cumulative Effects",
        x="lag", y="Autocorrelation") +
```



```

theme_bw()
ggsave("acf11.pdf", width=7, height=4)

## 10.3.3 Analysis 16
### Trace plot
pp <- do.call(cbind, joint.fit4.cum$mcmc["Dalphas"]) %>% data.frame()

ggplot(aes(x=seq(1, nrow(pp)), y=pp[,1]), data=pp) +
  geom_line() +
  labs(title=TeX("Trace Plot for Cumulative Association Structure  $\alpha^{(p)}_{(z)}$ "),
        subtitle="Analysis 16: Joint Model on CD4 Cell Counts and Viral Peak Point
                  with Shared Cumulative Effects",
        x="iteration", y=TeX(" $\alpha^{(p)}_{(z)}$ ")) +
  theme_bw()
ggsave("trace12.pdf", width=7, height=3.5)

### ACF plot
pp1 <- with(acf(pp[,1], plot=FALSE), data.frame(lag, acf))

ggplot(aes(x=lag, y=acf), data=pp1) +
  geom_hline(aes(yintercept=0)) +
  geom_segment(aes(xend=lag, yend=0)) +
  labs(title=TeX("Autocorrelation Plot for Cumulative Association Structure
                   $\alpha^{(p)}_{(z)}$ "),
        subtitle="Analysis 16: Joint Model on CD4 Cell Counts and Viral Peak Point
                  with Shared Cumulative Effects",
        x="lag", y="Autocorrelation") +
  theme_bw()
ggsave("acf12.pdf", width=7, height=4)

```

10.4 Sensitivity Analysis 7

```

## 10.4.1 RNA viral loads and time to rebound
joint.fit1.cum.sen0 <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                                     n.iter=50000, param="td-extra", extraForm=iForm,
                                     control=list(priorVar=1000))

summary(joint.fit1.cum.sen0)

## 10.4.2 RNA viral loads and time to peak point
joint.fit2.cum.sen0 <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                                     n.iter=50000, param="td-extra", extraForm=iForm,
                                     control=list(priorVar=1000))

summary(joint.fit2.cum.sen0)

## 10.4.3 CD4 cell counts and time to rebound
joint.fit3.cum.sen0 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                                     n.iter=50000, param="td-extra", extraForm=iForm,
                                     control=list(priorVar=1000))

summary(joint.fit3.cum.sen0)

## 10.4.4 CD4 cell counts and time to peak point
joint.fit4.cum.sen0 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",

```

```

n.iter=50000, param="td-extra", extraForm=iForm,
control=list(priorVar=1000))
summary(joint.fit4.cum.sen0)

```

10.5 Sensitivity Analysis 8

```

## 10.5.1 RNA viral loads and time to rebound
joint.fit1.cum.sen <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
n.iter=50000, param="td-extra", extraForm=iForm,
control=list(priorVar=1))

summary(joint.fit1.cum.sen)

## 10.5.2 RNA viral loads and time to peak point
joint.fit2.cum.sen <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
n.iter=50000, param="td-extra", extraForm=iForm,
control=list(priorVar=1))

summary(joint.fit2.cum.sen)

## 10.5.3 CD4 cell counts and time to rebound
joint.fit3.cum.sen <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
n.iter=50000, param="td-extra", extraForm=iForm,
control=list(priorVar=1))

summary(joint.fit3.cum.sen)

## 10.5.4 CD4 cell counts and time to peak point
joint.fit4.cum.sen <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
n.iter=50000, param="td-extra", extraForm=iForm,
control=list(priorVar=1))

summary(joint.fit4.cum.sen)

```

10.6 Sensitivity Analysis 9

```

## 10.6.1 RNA viral loads and time to rebound
joint.fit1.cum.sen2 <- jointModelBayes(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
n.iter=50000, param="td-extra", extraForm=iForm,
control=list(priorVar=0.1))

summary(joint.fit1.cum.sen2)

## 10.6.2 RNA viral loads and time to peak point
joint.fit2.cum.sen2 <- jointModelBayes(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
n.iter=50000, param="td-extra", extraForm=iForm,
control=list(priorVar=0.1))

summary(joint.fit2.cum.sen2)

## 10.6.3 CD4 cell counts and time to rebound
joint.fit3.cum.sen2 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
n.iter=50000, param="td-extra", extraForm=iForm,
control=list(priorVar=0.1))

summary(joint.fit3.cum.sen2)

```

```
## 10.6.4 CD4 cell counts and time to peak point
joint.fit4.cum.sen2 <- jointModelBayes(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                                     n.iter=50000, param="td-extra", extraForm=iForm,
                                     control=list(priorVar=0.1))

summary(joint.fit4.cum.sen2)
```

11. Individual Prediction

```
## 11.1 Survival prediction
nd.surv <- dat %>%
  mutate(t1=months_from_seroco-art_start) %>%
  filter(PATIENT==1 & treatment==0 & t1<23.73333)
sfit <- survfitJM(joint.fit1, newdata=nd.surv, idVar="PATIENT")

pdf("survpred.pdf", width=6, height=4)
plot(sfit, estimator="mean", include.y=TRUE, conf.int=TRUE, fill.area=TRUE, xlim=c(21,31),
     main="Survival Prediction for Patient 1", xlab="Month since the start of ART",
     ylab="Survival Probability",
     ylab2="RNA Viral Load (log10-scale)")
dev.off()

## 11.2 Longitudinal Prediction
nd.long <- dat %>%
  mutate(t1=months_from_seroco-art_start) %>%
  filter(PATIENT==1 & treatment==1)
lfit <- predict(joint.fit1, newdata=nd.long, idVar="PATIENT", type="Subject",
               interval="confidence", return=TRUE, FtTimes=seq(5.966667, 8, length.out=25))
last.time <- max(with(lfit, t1[is.na(low)]))

pdf("longpred.pdf", width=6, height=4)
xyplot(pred+low+upp~t1, data=lfit, type="l", lty=c(1,2,2),
       col=c(2,1,1), abline=list(v=last.time, lty=3),
       main="Longitudinal Prediction for Patient 1",
       xlab="Month since the start of ART",
       ylab=("Predicted RNA Viral Load (log10-scale)"))
dev.off()
```

12. Joint Models with Likelihood Method (shared $\eta_i(t)$)

```
## 12.1 RNA viral loads and time to rebound
jm.fit1.eta <- jointModel(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                         parameterization="value", method="spline-PH-GH")

summary(jm.fit1.eta)

## 12.2 RNA viral loads and time to peak point
jm.fit2.eta <- jointModel(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                         parameterization="value", method="spline-PH-GH")

summary(jm.fit2.eta)
```

```
## 12.3 CD4 cell counts and time to rebound
jm.fit3.eta <- jointModel(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                        parameterization="value", method="spline-PH-GH")
summary(jm.fit3.eta)

## 12.4 CD4 cell counts and time to peak point
jm.fit4.eta <- jointModel(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                        parameterization="value", method="spline-PH-GH")
summary(jm.fit4.eta)
```

13. Joint Model with Likelihood Method (shared $\eta_i(t)$ and $\eta'_i(t)$)

```
## 13.1 RNA viral loads and time to rebound
dForm <- list(fixed=~1, random=~1, indFixed=2, indRandom=2)
jm.fit1.eta.p <- jointModel(lme.fit.rna.rm, cox.fit.rebound, timeVar="t1",
                          parameterization="both",
                          method="piecewise-PH-GH", derivForm=dForm)
summary(jm.fit1.eta.p)

## 13.2 RNA viral loads and time to peak point
jm.fit2.eta.p <- jointModel(lme.fit.rna.rm, cox.fit.peak, timeVar="t1",
                          parameterization="both",
                          method="piecewise-PH-GH", derivForm=dForm)
summary(jm.fit2.eta.p)

## 13.3 CD4 cell counts and time to rebound
jm.fit3.eta.p <- jointModel(lme.fit.cd4.rm, cox.fit.rebound, timeVar="t1",
                          parameterization="both",
                          method="piecewise-PH-GH", derivForm=dForm)
summary(jm.fit3.eta.p)

## 13.4 CD4 cell counts and time to peak point
jm.fit4.eta.p <- jointModel(lme.fit.cd4.rm, cox.fit.peak, timeVar="t1",
                          parameterization="both",
                          method="piecewise-PH-GH", derivForm=dForm)
summary(jm.fit4.eta.p)
```

14. Model Comparison for Bayesian Joint Models

```
anova(joint.fit1, joint.fit1.eta, joint.fit1.eta.p, joint.fit1.cum)
anova(joint.fit2, joint.fit2.eta, joint.fit2.eta.p, joint.fit2.cum)
anova(joint.fit3, joint.fit3.eta, joint.fit3.eta.p, joint.fit3.cum)
anova(joint.fit4, joint.fit4.eta, joint.fit4.eta.p, joint.fit4.cum)
```