**Project:** Project 1, BIOS 6623

**Report:** Descriptive Statistics & Bayesian Analysis

**Investigator:** Rachel Johnson

**Date:** 9 October 2017

**Introduction**

The goals of this project are to examine the difference in treatment responses of hard drug users and those who don’t report hard drug use in HIV-1 infected homosexual and bisexual men in the Multicenter AIDS cohort study. Outcomes of interest include two laboratory measures—viral load and CD4+ T cell count—and two quality of life scores—aggregate mental score and aggregate physical score from the SF-36.

These measures were taken on men at baseline and throughout the follow-up period of the study, but this analysis aims to examine how these treatment responses differ between users and non-users of hard drugs from baseline values to their responses at the 2 year follow-up time point in the study.

**Methods**

There were 715 men included in the study at baseline, but 209 of these men did not have data at the 2 year follow-up, so they were excluded. Of the remaining 506, 28 were excluded since they did not have data available for all 4 outcomes, which was acceptable since none of those excluded were hard drug users, so it didn’t exclude any data on our population of interest.

The following demographic variables and potential covariates were selected in addition to hard drug use on the basis of covariate adjustment in previous studies: baseline outcome values, age, BMI, race, baseline marijuana use, baseline alcohol use, baseline smoking status, income, education, and adherence to ART at 2 years.

Some of these variables’ categories were collapsed to match how they were categorized in previous studies. Race was categorized into non-Hispanic white vs. other (reference); alcohol use was categorized into > 13 drinks per week vs. 13 or fewer drinks per week (reference); smoking was categorized into current smokers vs. never/former smokers (reference); income was categorized into < $10,000 (reference), $10,000 - $40,000, and > $40,000; education was categorized into greater than a high school education vs. a high school education or less (reference); and ART adherence at the 2 year mark was categorized into > 95% vs. < 95% (reference).

Continuous variables all remained continuous for the purposes of the analysis and descriptive statistics. Impossible values for BMI (> 500) were removed, and viral load was transformed on the log10 scale.

The linear regression models for the differences of each outcome as functions of all previously listed demographics variables and hard drug use were fitted in a Bayesian framework. All coefficients had vague priors that were normally distributed with means of 0 and variances of 1000, except for the parameters in the CD4+ count model, which had variances of 10000 under the normal distribution with mean 0. The variance had an inverse gamma prior with a shape of 2.001 and a scale of 1.001.

PROC MCMC was utilized to simulate the posterior distribution for each outcome with random walk Metropolis-Hastings algorithm. All covariate parameters had starting values of 0, and the error had a starting value of 1.

All models were run with a single chain, but each outcome required different numbers of iterations to reach convergence with acceptable mixing. Both quality of life outcomes were run for 30000 iterations with 2500 iterations discarded for burn-in. The log10 viral load models were run for 50000 iterations with 2500 burn-in iterations. Finally, the models for CD4+ count were run with 7500 burn-in iterations and 250000 total iterations, with only every 10th iteration included in the posterior distribution.

The Geweke statistic was used to assess model convergence, and model mixing was assessed with trace plots and autocorrelation plots. The results of these models were summarized with the posterior mean and 95% HPD credible intervals.

A model was fitted for each of the four difference outcomes with only hard drug use and baseline values for each outcome as the predictors. Then, full models were fitted for each of the difference outcomes, which included hard drug use, baseline outcome value, and the other demographics previously listed.

To explore which variables contributed most to model fit, one covariate was removed at a time from the full model, and the DIC from the full and reduced model were compared to determine if the model was improved with this variable included. It was requested that hard drug use, baseline outcome value, age, BMI, drug adherence, race, and at least one measure of socioeconomic status (education or income) remain in the model throughout the process. Therefore, models were tested with the removal of baseline values of alcohol consumption, smoking status, marijuana use, and income or education. Once all DIC values were compared, all variables that improved DIC when removed were all removed in one step, and this DIC was compared to that of the full model to determine if they could all be removed at once. If so, the process of removing the removable variables one at a time and comparing DIC to the new full model was repeated.

None of the estimates of hard drug use changed between the crude models and full model, so no further exploration into confounding or correlation was required to select final models.

**Results**

Table 1 reports the mean and standard deviation for continuous demographic variables and count and percentage for categorical demographics variables, and it is stratified by hard drug use.

Table 2 contains the differences from baseline to the 2 year mark in the study for each outcome, also stratified by hard drug use.

Table 3 includes the posterior means and the 95% HPD intervals for the hard drug use variable for each outcome’s model, adjusted for baseline, age, BMI, ART adherence at the 2 year mark, race, and income.

Hard drug use was not a strong predictor of the difference in SF-36 MCS over the course of 2 years after adjusting for age, BMI, race, income, adherence, and baseline values; hard drug users had a mean difference in SF-36 aggregate mental score that was 0.2167 points lower than non-hard drug users (95% HPD interval: -3.4630, 3.33566). Hard drug use was a much stronger predictor of the difference in SF-36 aggregate physical score after the same covariate adjustments. Hard drug users had a mean difference in physical quality of life score that was 3.6128 points lower than the physical score of non-hard drug users (95% HPD interval: -5.9669, 3.33566).

Hard drug use was also not a strong predictor of difference in log10 viral load after adjustment for age, BMI, race, income, adherence, and baseline values. **NEED TO SAY STUFF HERE.** However, hard drug use was a much stronger predictor of CD4+ count after these covariate adjustments. Hard drug users had a difference in CD4+ count that was 157.4 less than non-hard drug users (95% HPD interval: -216.4, -100.0).

**Table 1.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Total** | **Hard drugs = Yes** | **Hard drugs = No** |
| **Age at baseline** | 43.26 ± 8.72 | 44.62 ± 9.49 | 43.14 ± 8.65 |
| **BMI at baseline** | 25.21 ± 4.34 | 23.62 ± 3.45 | 25.34 ± 4.39 |
| **Alcohol use at baseline** |  |  |  |
| 13 or fewer drinks per week | 443 (93.07) | 37 (94.87) | 406 (92.91) |
| > 13 drinks per week | 33 (6.93) | 2 (5.13) | 31 (7.09) |
| **Smoking status at baseline** |  |  |  |
| Never/former | 291 (61.13) | 9 (23.08) | 282 (64.53) |
| Current | 185 (38.87) | 30 (76.92) | 155 (35.47) |
| **Marijuana use at baseline** |  |  |  |
| Yes | 194 (40.76) | 12 (30.77) | 182 (41.65) |
| No | 282 (59.24) | 27 (69.23) | 255 (58.35) |
| **Income level at baseline** |  |  |  |
| < $10,000 | 99 (20.8) | 14 (35.9) | 85 (20.24) |
| $10,000 - $40,000 | 196 (41.18) | 16 (41.03) | 180 (42.86) |
| > $40,000 | 164 (34.45) | 9 (23.08) | 155 (36.9) |
| **Education at baseline** |  |  |  |
| HS or less | 101 (21.22) | 16 (41.03) | 85 (19.45) |
| >HS | 375 (78.78) | 23 (58.97) | 352 (80.55) |
| **Adherence at 2 years** |  |  |  |
| <95% | 50 (10.5) | 1 (2.56) | 49 (11.21) |
| >95% | 426 (89.5) | 38 (97.44) | 388 (88.79) |
| **Baseline log10 viral load** | 4.52 ± 0.92 | 4.52 ± 0.86 | 4.52 ± 0.93 |
| **Baseline CD4+ count** | 373.53 ± 200.46 | 352.18 ± 194.67 | 375.44 ± 201.07 |
| **Baseline SF36 MCS score** | 44.89 ± 13.5 | 42.31 ± 11.22 | 45.12 ± 13.68 |
| **Baseline SF36 PCS score** | 51.04 ± 9.05 | 47.7 ± 8.5 | 51.34 ± 9.05 |

**Table 2.**  This table displays the difference in the four outcomes from baseline to the 2 year mark in the study.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Total** | **Hard drugs = Yes** | **Hard drugs = No** |
| **Difference in log10 viral load** | 0.41 ± 0.37 | 0.39 ± 0.25 | 0.41 ± 0.38 |
| **Difference in CD4+ count** | 169.21 ± 183.2 | 13.43 ± 195.73 | 183.11 ± 175.67 |
| **Difference in SF36 MCS score** | 2.34 ± 12.1 | 3.58 ± 15.07 | 2.23 ± 11.82 |
| **Difference in SF36 PCS score** | -1.55 ± 8.37 | -3.85 ± 8.71 | -1.34 ± 8.32 |

**Table 3.** This table shows the posterior means and 95% HPD intervals for the hard drug use variable in each of the 4 covariate-adjusted models.

|  |  |  |
| --- | --- | --- |
| **Outcome** | **Hard drug posterior mean** | **Hard drug 95% HPD interval** |
| SF-36 MCS (Mental)\* | -0.2167 | -3.4630, 3.33566 |
| SF-36 PCS (Physical)\* | -3.6128 | -5.9669, -0.8246 |
| Log10 Viral Load\* | -0.0363 | -0.1595, 0.1015 |
| CD4+ Count\* | -157.4 | -216.4, -100.0 |

*\*Adjusted for baseline, age, BMI, ART adherence at 2 year mark, race, income*

**Table 4.** This table displays the results of the final model for the difference in SF-36 MCS score from baseline to the 2 year mark.

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Mean** | **95% HPD Interval** | |
| **Intercept** | 19.4077 | 11.5768 | 26.6034 |
| **Baseline SF-36 MCS** | -0.5363 | -0.6096 | -0.4696 |
| **Hard drug use** | -0.2167 | -3.463 | 3.3566 |
| **Age** | 0.091 | -0.0254 | 0.1973 |
| **BMI** | 0.00921 | -0.1952 | 0.2233 |
| **Adherence** | 2.5114 | -0.6865 | 5.4813 |
| **Race** | -0.2294 | -2.331 | 2.0777 |
| **Income (Medium)** | 0.6213 | -1.8418 | 3.1895 |
| **Income (High)** | 2.0207 | -0.7472 | 4.8331 |
| **Variance** | 98.797 | 86.01 | 111.9 |

**Table 5.** This table displays the results of the final model for the difference in SF-36 PCS score from baseline to the 2 year mark.

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Mean** | **95% HPD Interval** | |
| **Intercept** | 14.7989 | 7.59 | 22.1062 |
| **Baseline SF-36 PCS** | -0.3152 | -0.4026 | -0.2303 |
| **Hard drug use** | -3.6128 | -5.9669 | -0.8246 |
| **Age** | -0.1158 | -0.198 | -0.0262 |
| **BMI** | 0.0746 | -0.1001 | 0.2433 |
| **Adherence** | 1.5434 | -0.9391 | 3.9661 |
| **Race** | -1.0081 | -2.6029 | 0.6921 |
| **Income (Medium)** | 2.58 | 0.7087 | 4.4375 |
| **Income (High)** | 3.5038 | 1.3955 | 5.7371 |
| **Variance** | 58.0263 | 50.3957 | 65.7037 |

**Table 6.** This table displays the results of the final model for the difference in log10 viral load score from baseline to the 2 year mark.

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Mean** | **95% HPD Interval** | |
| **Intercept** | 1.2247 | 0.8845 | 1.5665 |
| **Baseline log10 viral load** | -0.0854 | -0.1242 | -0.0475 |
| **Hard drug use** | -0.0363 | -0.1595 | 0.1015 |
| **Age** | -0.0005 | -0.0045 | 0.00326 |
| **BMI** | -0.0074 | -0.0154 | 0.00066 |
| **Adherence** | -0.1442 | -0.261 | -0.0283 |
| **Race** | -0.0036 | -0.085 | 0.0775 |
| **Income (Medium)** | -0.0682 | -0.1621 | 0.0263 |
| **Income (High)** | -0.1508 | -0.2528 | -0.0433 |
| **Variance** | 0.1344 | 0.1172 | 0.1526 |

**Table 7.** This table displays the results of the final model for the difference in CD4+ count from baseline to the 2 year mark.

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Mean** | **95% HPD Interval** | |
| **Intercept** | 43.1248 | -64.414 | 143.9 |
| **Baseline CD4+ count** | -0.0463 | -0.1241 | 0.0332 |
| **Hard drug use** | -157.4 | -216.4 | -100 |
| **Age** | -0.3321 | -2.0326 | 1.383 |
| **BMI** | 4.5319 | 0.9319 | 8.5578 |
| **Adherence** | 57.8586 | 7.7837 | 107.9 |
| **Race** | 28.5692 | -9.7254 | 63.5578 |
| **Income (medium)** | 5.3645 | -37.778 | 47.1722 |
| **Income (high)** | -48.476 | -94.295 | -2.9473 |
| **Variance** | 29902.9 | 26062 | 34039.9 |

**Conclusion**

Overall, there were increases in the mean differences in CD4+ count and SF-36 MCS, and there was an overall decrease in mean difference in SF-36 PCS (Table 3). This supports the biological explanation that CD4+ counts should recover after treatment is initiated. It also supports the idea that treatment can help improve quality of life, but that the side effects may be physically intense.

After covariate adjustment, hard drug use was a strong predictor of both changes in aggregate physical score and CD4+ count. Biologically, this supports the limited laboratory evidence that hard drug use inhibits the immune system and therefore slows the recovery of CD4+ count. Greater decreases in aggregate score in hard drug users also make medical sense: greater potential inhibition of the immune system and the overall health damage due to hard drug use could contribute to a lower physical quality of life.

The study is potentially limited by the small sample size of hard drug users at baseline (n = 39). Furthermore, there are only four major cities involved in the study. It may have been helpful to account for differences between cities or regions, but based on the small sample size of the hard drug user group, it would be difficult to stratify the population any further.

https://github.com/BIOS6623-UCD/bios6623-johnsra3/tree/master/Project1

**Code**

#=============================================================#

# Import data

#=============================================================#

library(dplyr)

setwd("~/School/AdvancedData")

hiv\_initial <- read.csv("hiv\_6623\_final\_nodots.csv", header = T)

hiv0 <- hiv\_initial[hiv\_initial$years == 0, ] #715 people at baseline

hiv2 <- hiv\_initial[hiv\_initial$years == 2, ] #506 people at 2 yr mark

#=============================================================#

# Clean data

#=============================================================#

##################

#Remove people who don't have data at 2 years

hiv0 <- hiv0[hiv0$newid %in% hiv2$newid, ]

hiv0 <- hiv0[, c(which(colnames(hiv0) == "newid"),

which(colnames(hiv0) == "AGG\_MENT"),

which(colnames(hiv0) == "AGG\_PHYS"),

which(colnames(hiv0) == "LEU3N"),

which(colnames(hiv0) == "VLOAD"),

which(colnames(hiv0) == "hard\_drugs"),

which(colnames(hiv0) == "age"),

which(colnames(hiv0) == "BMI"),

which(colnames(hiv0) == "RACE"),

which(colnames(hiv0) == "HASHV"),

which(colnames(hiv0) == "DKGRP"),

which(colnames(hiv0) == "SMOKE"),

which(colnames(hiv0) == "income"),

which(colnames(hiv0) == "EDUCBAS"))]

hiv2 <- hiv2[, c(which(colnames(hiv2) == "newid"),

which(colnames(hiv2) == "ADH"),

which(colnames(hiv2) == "AGG\_MENT"),

which(colnames(hiv2) == "AGG\_PHYS"),

which(colnames(hiv2) == "LEU3N"),

which(colnames(hiv2) == "VLOAD"))]

###################

#Need to categorize race, alcohol use, smoke, income, education, ART adherence

#Race- change to NHW (non-hispanic white) vs. Other

table(hiv0$RACE)

hiv0$RACE\_cat <- hiv0$RACE

hiv0$RACE\_cat <- ifelse(hiv0$RACE\_cat == 1, "Non-Hispanic White", "Other")

table(hiv0$RACE\_cat)

#Alcohol

table(hiv0$DKGRP)

hiv0$drink\_cat <- hiv0$DKGRP

hiv0$drink\_cat <- ifelse(hiv0$DKGRP == 3, "> 13 drinks per week", "13 or fewer drinks per week")

table(hiv0$drink\_cat)

#Marijuana

table(hiv0$HASHV)

hiv0$marij\_cat <- hiv0$HASHV

hiv0$marij\_cat <- ifelse(hiv0$HASHV == 2, "Yes", "No")

table(hiv0$marij\_cat)

#Smoking

table(hiv0$SMOKE)

hiv0$smoke\_cat <- hiv0$SMOKE

hiv0$smoke\_cat <- ifelse(hiv0$smoke\_cat == 3, "Current", "Never/former")

table(hiv0$smoke\_cat)

#Income

table(hiv0$income)

hiv0$income\_cat <- hiv0$income

hiv0$income\_cat[hiv0$income\_cat == 9] <- NA #remove those who chose not to answer

hiv0$income\_cat <- as.character(hiv0$income\_cat)

hiv0$income\_cat[hiv0$income\_cat == "1"] <- "< $10,000"

hiv0$income\_cat[hiv0$income\_cat == "2"] <- "$10,000 - $40,000"

hiv0$income\_cat[hiv0$income\_cat == "3"] <- "$10,000 - $40,000"

hiv0$income\_cat[hiv0$income\_cat == "4"] <- "$10,000 - $40,000"

hiv0$income\_cat[hiv0$income\_cat == "5"] <- "> $40,000"

hiv0$income\_cat[hiv0$income\_cat == "6"] <- "> $40,000"

hiv0$income\_cat[hiv0$income\_cat == "7"] <- "> $40,000"

#Education

table(hiv0$EDUCBAS)

hiv0$educ\_cat <- hiv0$EDUCBAS

hiv0$educ\_cat <- ifelse(hiv0$educ\_cat > 3, ">HS", "HS or less")

table(hiv0$educ\_cat)

# ART Adherence, >95% vs. <95%

table(hiv2$ADH)

hiv2$adh\_cat <- hiv2$ADH

hiv2$adh\_cat <- ifelse(hiv2$adh\_cat < 3, ">95%", "<95%")

table(hiv2$adh\_cat)

####################

#Format categorical variables for table

names(hiv0)

#hard\_drugs

table(hiv0$hard\_drugs)

hiv0$hard\_drugs <- factor(hiv0$hard\_drugs, levels = c("0", "1"),

labels = c("No", "Yes"))

#marij\_cat

table(hiv0$marij\_cat)

hiv0$marij\_cat <- factor(hiv0$marij\_cat, levels = c("Yes", "No"))

#RACE\_cat

table(hiv0$RACE\_cat)

hiv0$RACE\_cat <- factor(hiv0$RACE\_cat, levels = c("Non-Hispanic White", "Other"))

#drink\_cat

table(hiv0$drink\_cat)

hiv0$drink\_cat <- factor(hiv0$drink\_cat, levels = c("13 or fewer drinks per week", "> 13 drinks per week"))

#smoke\_cat

table(hiv0$smoke\_cat)

hiv0$smoke\_cat <- factor(hiv0$smoke\_cat, levels = c("Never/former", "Current"))

#income\_cat

table(hiv0$income\_cat)

hiv0$income\_cat <- factor(hiv0$income\_cat, levels = c("< $10,000", "$10,000 - $40,000", "> $40,000"))

#educ\_cat

table(hiv0$educ\_cat)

hiv0$educ\_cat <- factor(hiv0$educ\_cat, levels = c("HS or less", ">HS"))

#adh\_cat (at 2 years)

table(hiv2$adh\_cat)

hiv2$adh\_cat <- factor(hiv2$adh\_cat, levels = c("<95%", ">95%"))

#=============================================================#

# Merge datasets and create difference variables

#=============================================================#

####Merge

names(hiv0)

names(hiv2)

hiv0\_merge <- hiv0[, c(which(colnames(hiv0) == "newid"),

which(colnames(hiv0) == "AGG\_MENT"),

which(colnames(hiv0) == "AGG\_PHYS"),

which(colnames(hiv0) == "LEU3N"),

which(colnames(hiv0) == "VLOAD"),

which(colnames(hiv0) == "hard\_drugs"),

which(colnames(hiv0) == "age"),

which(colnames(hiv0) == "BMI"),

which(colnames(hiv0) == "RACE\_cat"),

which(colnames(hiv0) == "drink\_cat"),

which(colnames(hiv0) == "marij\_cat"),

which(colnames(hiv0) == "smoke\_cat"),

which(colnames(hiv0) == "income\_cat"),

which(colnames(hiv0) == "educ\_cat"))]

colnames(hiv2)[3:7] <- c("AGG\_MENT\_2yr", "AGG\_PHYS\_2yr", "LEU3N\_2yr",

"VLOAD\_2yr", "adh\_cat\_2yr")

hiv <- merge(hiv0\_merge, hiv2, by = "newid")

###########

# Create difference variables

hiv$diff\_leu3n <- hiv$LEU3N\_2yr - hiv$LEU3N

hiv$diff\_aggment <- hiv$AGG\_MENT\_2yr - hiv$AGG\_MENT

hiv$diff\_aggphys <- hiv$AGG\_PHYS\_2yr - hiv$AGG\_PHYS

#Log10 transform vload and then find difference

hiv$logvload <- log10(hiv$VLOAD)

hiv$logvload\_2yr <- log10(hiv$VLOAD\_2yr)

boxplot(hiv$logvload, horizontal = TRUE)

boxplot(hiv$logvload\_2yr, horizontal = TRUE)

hiv$diff\_logvload <- hiv$logvload\_2yr/hiv$logvload

##########

test <- hiv

table(test$hard\_drugs) #467 no, 39 yes

test <- test[is.na(test$diff\_aggment) == FALSE, ]

test <- test[is.na(test$diff\_aggphys) == FALSE, ]

test <- test[is.na(test$diff\_leu3n) == FALSE, ]

test <- test[is.na(test$diff\_logvload) == FALSE, ]

table(test$hard\_drugs) #439 no, 39 yes-- doesn't affect # of ppl who yes hard drugs

hiv <- test

extra <- hiv

#=============================================================#

# Summarize rest of cont. variables to get rid of NA coding

#=============================================================#

summary(hiv$BMI)

hiv$BMI[hiv$BMI == -1] <- NA

hiv$BMI[hiv$BMI == 999] <- NA

#still have a maximum of 514.25-- changing to NA for data entry

hiv$BMI[hiv$BMI > 500] <- NA

#all seem probable, even if some are fairly low or high

boxplot(hiv$BMI, horizontal = TRUE)

summary(hiv$age)

#all seem probable

#=============================================================#

# Create table 1: demographics

#=============================================================#

############

# Make dataset for drugyes and drugno

table(hiv$hard\_drugs)

hivyes <- hiv[hiv$hard\_drugs == "Yes", ]

hivno <- hiv[hiv$hard\_drugs == "No", ]

#need rows for age (1), bmi (1), hard\_drugs (3), race\_cat (3),

#drink\_cat (3), smoke\_cat (3), marij\_cat (3), income\_cat (4), educ\_cat (3),

#ADH (3), baseline vload (1), baseline leu3n (1),

#baseline aggment (1), baseline aggphys (1)

demtab <- matrix(data = NA, nrow = 25, ncol = 4)

colnames(demtab) <- c("", "Total", "Hard drugs = Yes", "Hard drugs = No")

demtab[1, 1] <- "Age at baseline"

demtab[1, 2] <- paste(round(mean(hiv$age), 2), "±", round(sd(hiv$age), 2))

demtab[1, 3] <- paste(round(mean(hivyes$age), 2), "±", round(sd(hivyes$age), 2))

demtab[1, 4] <- paste(round(mean(hivno$age), 2), "±", round(sd(hivno$age), 2))

demtab[2, 1] <- "BMI at baseline"

demtab[2, 2] <- paste(round(mean(hiv$BMI, na.rm = T), 2), "±", round(sd(hiv$BMI, na.rm = T), 2))

demtab[2, 3] <- paste(round(mean(hivyes$BMI, na.rm = T), 2), "±", round(sd(hivyes$BMI, na.rm = T), 2))

demtab[2, 4] <- paste(round(mean(hivno$BMI, na.rm = T), 2), "±", round(sd(hivno$BMI, na.rm = T), 2))

demtab[3, 1] <- "Alcohol use at baseline"

demtab[4:5, 1] <- levels(hiv$drink\_cat)

demtab[4, 2] <- paste(nrow(hiv[hiv$drink\_cat == "13 or fewer drinks per week", ]),

paste("(", round(nrow(hiv[hiv$drink\_cat == "13 or fewer drinks per week", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[5, 2] <- paste(nrow(hiv[hiv$drink\_cat == "> 13 drinks per week", ]),

paste("(", round(nrow(hiv[hiv$drink\_cat == "> 13 drinks per week", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[4, 3] <- paste(nrow(hivyes[hivyes$drink\_cat == "13 or fewer drinks per week", ]),

paste("(", round(nrow(hivyes[hivyes$drink\_cat == "13 or fewer drinks per week", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[5, 3] <- paste(nrow(hivyes[hivyes$drink\_cat == "> 13 drinks per week", ]),

paste("(", round(nrow(hivyes[hivyes$drink\_cat == "> 13 drinks per week", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[4, 4] <- paste(nrow(hivno[hivno$drink\_cat == "13 or fewer drinks per week", ]),

paste("(", round(nrow(hivno[hivno$drink\_cat == "13 or fewer drinks per week", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[5, 4] <- paste(nrow(hivno[hivno$drink\_cat == "> 13 drinks per week", ]),

paste("(", round(nrow(hivno[hivno$drink\_cat == "> 13 drinks per week", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[6, 1] <- "Smoking status at baseline"

demtab[7:8, 1] <- levels(hiv$smoke\_cat)

demtab[7, 2] <- paste(nrow(hiv[hiv$smoke\_cat == "Never/former", ]),

paste("(", round(nrow(hiv[hiv$smoke\_cat == "Never/former", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[8, 2] <- paste(nrow(hiv[hiv$smoke\_cat == "Current", ]),

paste("(", round(nrow(hiv[hiv$smoke\_cat == "Current", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[7, 3] <- paste(nrow(hivyes[hivyes$smoke\_cat == "Never/former", ]),

paste("(", round(nrow(hivyes[hivyes$smoke\_cat == "Never/former", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[8, 3] <- paste(nrow(hivyes[hivyes$smoke\_cat == "Current", ]),

paste("(", round(nrow(hivyes[hivyes$smoke\_cat == "Current", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[7, 4] <- paste(nrow(hivno[hivno$smoke\_cat == "Never/former", ]),

paste("(", round(nrow(hivno[hivno$smoke\_cat == "Never/former", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[8, 4] <- paste(nrow(hivno[hivno$smoke\_cat == "Current", ]),

paste("(", round(nrow(hivno[hivno$smoke\_cat == "Current", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[9, 1] <- "Marijuana use at baseline"

demtab[10:11, 1] <- levels(hiv$marij\_cat)

demtab[10, 2] <- paste(nrow(hiv[hiv$marij\_cat == "Yes", ]),

paste("(", round(nrow(hiv[hiv$marij\_cat == "Yes", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[10, 3] <- paste(nrow(hivyes[hivyes$marij\_cat == "Yes", ]),

paste("(", round(nrow(hivyes[hivyes$marij\_cat == "Yes", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[10, 4] <- paste(nrow(hivno[hivno$marij\_cat == "Yes", ]),

paste("(", round(nrow(hivno[hivno$marij\_cat == "Yes", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[11, 2] <- paste(nrow(hiv[hiv$marij\_cat == "No", ]),

paste("(", round(nrow(hiv[hiv$marij\_cat == "No", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[11, 3] <- paste(nrow(hivyes[hivyes$marij\_cat == "No", ]),

paste("(", round(nrow(hivyes[hivyes$marij\_cat == "No", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[11, 4] <- paste(nrow(hivno[hivno$marij\_cat == "No", ]),

paste("(", round(nrow(hivno[hivno$marij\_cat == "No", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[12, 1] <- "Income level at baseline"

demtab[13:15, 1] <- levels(hiv$income\_cat)

inc <- cbind.data.frame(hiv$income\_cat, hiv$newid, hiv$hard\_drugs)

inc <- inc[is.na(inc$`hiv$income\_cat`) == FALSE, ]

colnames(inc) <- c("income\_cat", "newid", "hard\_drugs")

incyes <- inc[inc$hard\_drugs == "Yes", ]

incno <- inc[inc$hard\_drugs == "No", ]

demtab[13, 2] <- paste(nrow(inc[inc$income\_cat == "< $10,000", ]),

paste("(", round(nrow(inc[inc$income\_cat == "< $10,000", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[14, 2] <- paste(nrow(inc[inc$income\_cat == "$10,000 - $40,000", ]),

paste("(", round(nrow(inc[inc$income\_cat == "$10,000 - $40,000", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[15, 2] <- paste(nrow(inc[inc$income\_cat == "> $40,000", ]),

paste("(", round(nrow(inc[inc$income\_cat == "> $40,000", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[13, 3] <- paste(nrow(incyes[incyes$income\_cat == "< $10,000", ]),

paste("(", round(nrow(incyes[incyes$income\_cat == "< $10,000", ])/nrow(incyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[14, 3] <- paste(nrow(incyes[incyes$income\_cat == "$10,000 - $40,000", ]),

paste("(", round(nrow(incyes[incyes$income\_cat == "$10,000 - $40,000", ])/nrow(incyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[15, 3] <- paste(nrow(incyes[incyes$income\_cat == "> $40,000", ]),

paste("(", round(nrow(incyes[incyes$income\_cat == "> $40,000", ])/nrow(incyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[13, 4] <- paste(nrow(incno[incno$income\_cat == "< $10,000", ]),

paste("(", round(nrow(incno[incno$income\_cat == "< $10,000", ])/nrow(incno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[14, 4] <- paste(nrow(incno[incno$income\_cat == "$10,000 - $40,000", ]),

paste("(", round(nrow(incno[incno$income\_cat == "$10,000 - $40,000", ])/nrow(incno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[15, 4] <- paste(nrow(incno[incno$income\_cat == "> $40,000", ]),

paste("(", round(nrow(incno[incno$income\_cat == "> $40,000", ])/nrow(incno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[16, 1] <- "Education at baseline"

demtab[17:18, 1] <- levels(hiv$educ\_cat)

demtab[17, 2] <- paste(nrow(hiv[hiv$educ\_cat == "HS or less", ]),

paste("(", round(nrow(hiv[hiv$educ\_cat == "HS or less", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[18, 2] <- paste(nrow(hiv[hiv$educ\_cat == ">HS", ]),

paste("(", round(nrow(hiv[hiv$educ\_cat == ">HS", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[17, 3] <- paste(nrow(hivyes[hivyes$educ\_cat == "HS or less", ]),

paste("(", round(nrow(hivyes[hivyes$educ\_cat == "HS or less", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[18, 3] <- paste(nrow(hivyes[hivyes$educ\_cat == ">HS", ]),

paste("(", round(nrow(hivyes[hivyes$educ\_cat == ">HS", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[17, 4] <- paste(nrow(hivno[hivno$educ\_cat == "HS or less", ]),

paste("(", round(nrow(hivno[hivno$educ\_cat == "HS or less", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[18, 4] <- paste(nrow(hivno[hivno$educ\_cat == ">HS", ]),

paste("(", round(nrow(hivno[hivno$educ\_cat == ">HS", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[19, 1] <- "Adherence at 2 years"

demtab[20:21, 1] <- levels(hiv$adh\_cat\_2yr)

demtab[20, 2] <- paste(nrow(hiv[hiv$adh\_cat\_2yr == "<95%", ]),

paste("(", round(nrow(hiv[hiv$adh\_cat\_2yr == "<95%", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[21, 2] <-paste(nrow(hiv[hiv$adh\_cat\_2yr == ">95%", ]),

paste("(", round(nrow(hiv[hiv$adh\_cat\_2yr == ">95%", ])/nrow(hiv) \* 100, 2), ")", sep = ""), sep = " ")

demtab[20, 3] <- paste(nrow(hivyes[hivyes$adh\_cat\_2yr == "<95%", ]),

paste("(", round(nrow(hivyes[hivyes$adh\_cat\_2yr == "<95%", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[21, 3] <-paste(nrow(hivyes[hivyes$adh\_cat\_2yr == ">95%", ]),

paste("(", round(nrow(hivyes[hivyes$adh\_cat\_2yr == ">95%", ])/nrow(hivyes) \* 100, 2), ")", sep = ""), sep = " ")

demtab[20, 4] <- paste(nrow(hivno[hivno$adh\_cat\_2yr == "<95%", ]),

paste("(", round(nrow(hivno[hivno$adh\_cat\_2yr == "<95%", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[21, 4] <- paste(nrow(hivno[hivno$adh\_cat\_2yr == ">95%", ]),

paste("(", round(nrow(hivno[hivno$adh\_cat\_2yr == ">95%", ])/nrow(hivno) \* 100, 2), ")", sep = ""), sep = " ")

demtab[22, 1] <- "Baseline log10 viral load"

demtab[22, 2] <- paste(round(mean(hiv$logvload), 2), "±", round(sd(hiv$logvload), 2))

demtab[22, 3] <- paste(round(mean(hivyes$logvload), 2), "±", round(sd(hivyes$logvload), 2))

demtab[22, 4] <- paste(round(mean(hivno$logvload), 2), "±", round(sd(hivno$logvload), 2))

demtab[23, 1] <- "Baseline CD4+ count"

demtab[23, 2] <- paste(round(mean(hiv$LEU3N), 2), "±", round(sd(hiv$LEU3N), 2))

demtab[23, 3] <- paste(round(mean(hivyes$LEU3N), 2), "±", round(sd(hivyes$LEU3N), 2))

demtab[23, 4] <- paste(round(mean(hivno$LEU3N), 2), "±", round(sd(hivno$LEU3N), 2))

demtab[24, 1] <- "Baseline SF36 MCS score"

demtab[24, 2] <- paste(round(mean(hiv$AGG\_MENT), 2), "±", round(sd(hiv$AGG\_MENT), 2))

demtab[24, 3] <- paste(round(mean(hivyes$AGG\_MENT), 2), "±", round(sd(hivyes$AGG\_MENT), 2))

demtab[24, 4] <- paste(round(mean(hivno$AGG\_MENT), 2), "±", round(sd(hivno$AGG\_MENT), 2))

demtab[25, 1] <- "Baseline SF36 PCS score"

demtab[25, 2] <- paste(round(mean(hiv$AGG\_PHYS), 2), "±", round(sd(hiv$AGG\_PHYS), 2))

demtab[25, 3] <- paste(round(mean(hivyes$AGG\_PHYS), 2), "±", round(sd(hivyes$AGG\_PHYS), 2))

demtab[25, 4] <- paste(round(mean(hivno$AGG\_PHYS), 2), "±", round(sd(hivno$AGG\_PHYS), 2))

#setwd("C:/Repositories/bios6623-johnsra3/Project1/Reports")

#write.csv(demtab, "DemographicsTable10022017.csv")

#=============================================================#

# Create table 2: outcomes

#=============================================================#

outtab <- matrix(data = NA, nrow = 4, ncol = 4)

colnames(outtab) <- c("", "Total", "Hard drugs = Yes", "Hard drugs = No")

outtab[1, 1] <- "Difference in log10 viral load"

outtab[1, 2] <- paste(round(mean(hiv$diff\_logvload, na.rm = T), 2),

"±", round(sd(hiv$diff\_logvload, na.rm = T), 2))

outtab[1, 3] <- paste(round(mean(hivyes$diff\_logvload, na.rm = T), 2),

"±", round(sd(hivyes$diff\_logvload, na.rm = T), 2))

outtab[1, 4] <- paste(round(mean(hivno$diff\_logvload, na.rm = T), 2),

"±", round(sd(hivno$diff\_logvload, na.rm = T), 2))

outtab[2, 1] <- "Difference in CD4+ count"

outtab[2, 2] <- paste(round(mean(hiv$diff\_leu3n, na.rm = T), 2),

"±", round(sd(hiv$diff\_leu3n, na.rm = T), 2))

outtab[2, 3] <- paste(round(mean(hivyes$diff\_leu3n, na.rm = T), 2),

"±", round(sd(hivyes$diff\_leu3n, na.rm = T), 2))

outtab[2, 4] <- paste(round(mean(hivno$diff\_leu3n, na.rm = T), 2),

"±", round(sd(hivno$diff\_leu3n, na.rm = T), 2))

outtab[3, 1] <- "Difference in SF36 MCS score"

outtab[3, 2] <- paste(round(mean(hiv$diff\_aggment, na.rm = T), 2), "±",

round(sd(hiv$diff\_aggment, na.rm = T), 2))

outtab[3, 3] <- paste(round(mean(hivyes$diff\_aggment, na.rm = T), 2), "±",

round(sd(hivyes$diff\_aggment, na.rm = T), 2))

outtab[3, 4] <- paste(round(mean(hivno$diff\_aggment, na.rm = T), 2), "±",

round(sd(hivno$diff\_aggment, na.rm = T), 2))

outtab[4, 1] <- "Difference in SF36 PCS score"

outtab[4, 2] <- paste(round(mean(hiv$diff\_aggphys, na.rm = T), 2), "±",

round(sd(hiv$diff\_aggphys, na.rm = T), 2))

outtab[4, 3] <- paste(round(mean(hivyes$diff\_aggphys, na.rm = T), 2), "±",

round(sd(hivyes$diff\_aggphys, na.rm = T), 2))

outtab[4, 4] <- paste(round(mean(hivno$diff\_aggphys, na.rm = T), 2), "±",

round(sd(hivno$diff\_aggphys, na.rm = T), 2))

setwd("C:/Repositories/bios6623-johnsra3/Project1/Reports")

write.csv(outtab, "DifferenceOutcomesTable09272017.csv")

#=============================================================#

# Create graphs for outcome differences

#=============================================================#

par(mfrow = c(2, 2))

boxplot(hiv$diff\_logvload ~ hiv$hard\_drugs, ylab = "Difference in Viral Load",

main = "Difference in log10 Viral Load \nby Hard Drug Use")

boxplot(hiv$diff\_leu3n ~ hiv$hard\_drugs, ylab = "Difference in CD4+ Count",

main = "Difference in CD4+ Count \nby Hard Drug Use")

boxplot(hiv$diff\_aggment ~ hiv$hard\_drugs, ylab = "Difference in SF36 MCS Score",

main = "Difference in SF36 MCS Score \nby Hard Drug Use")

boxplot(hiv$diff\_aggphys ~ hiv$hard\_drugs, ylab = "Difference in SF36 PCS Score",

main = "Difference in SF36 PCS Score \nby Hard Drug Use")

#=============================================================#

# Make dataset of indicator variables to import into SAS

#=============================================================#

names(extra)

names(hiv)

indic <- cbind.data.frame(extra, hiv[, 20:25])

names(indic)

indic <- indic[, -c(which(colnames(indic) == "AGG\_MENT\_2yr"),

which(colnames(indic) == "AGG\_PHYS\_2yr"),

which(colnames(indic) == "LEU3N\_2yr"),

which(colnames(indic) == "VLOAD\_2yr"),

which(colnames(indic) == "ADH"))]

names(indic)

summary(indic$hard\_drugs)

indic$harddrugsY <- NA

indic$harddrugsY <- ifelse(indic$hard\_drugs == "Yes", 1, 0)

indic <- indic[, -which(colnames(indic) == "hard\_drugs")]

indic$raceNHW <- NA

indic$raceNHW <- ifelse(indic$RACE\_cat == "Non-Hispanic White", 1, 0)

indic <- indic[, -which(colnames(indic) == "RACE\_cat")]

indic$drink13plus <- NA

indic$drink13plus <- ifelse(indic$drink\_cat == "> 13 drinks per week", 1, 0)

indic <- indic[, -which(colnames(indic) == "drink\_cat")]

indic$smokecurrent <- NA

indic$smokecurrent <- ifelse(indic$smoke\_cat == "Current", 1, 0)

indic <- indic[, -which(colnames(indic) == "smoke\_cat")]

indic$marijY <- NA

indic$marijY <- ifelse(indic$marij\_cat == "Yes", 1, 0)

indic <- indic[, -which(colnames(indic) == "marij\_cat")]

indic$incomehigh <- NA

indic$incomehigh <- ifelse(indic$income\_cat == "> $40,000", 1, 0)

indic$incomemed <- NA

indic$incomemed <- ifelse(indic$income\_cat == "$10,000 - $40,000", 1, 0)

indic <- indic[, -which(colnames(indic) == "income\_cat")]

indic$educHSmore <- NA

indic$educHSmore <- ifelse(indic$educ\_cat == ">HS", 1, 0)

indic <- indic[, -which(colnames(indic) == "educ\_cat")]

indic$adhhigh <- NA

indic$adhhigh <- ifelse(indic$adh\_cat\_2yr == ">95%", 1, 0)

indic <- indic[, -which(colnames(indic) == "adh\_cat\_2yr")]

setwd("~/School/AdvancedData")

write.csv(indic, "IndicatorHIVdata10022017.csv")

\*-----------------------------------------------------------;

\* Import indicator data frame from R code

\*-----------------------------------------------------------;

**PROC** **IMPORT** datafile = "C:\Users\johnsra3\Documents\School\AdvancedData\IndicatorHIVdata10022017\_dots.csv"

dbms = csv

out = hiv replace;

getnames = yes;

**RUN**;

\*-----------------------------------------------------------;

\* Crude model for AGG\_MENT (drugs and base only)

\*-----------------------------------------------------------;

\*DIC: 3547.111, better in big model, wh/ is good;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betaAM **0** betadrugs **0**;

PARMS sigma2 **1**;

PRIOR betaAM betadrugs ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAM\*AGG\_MENT + betadrugs\*harddrugsY;

model diff\_aggment ~ normal(mu, var = sigma2);

title "Crude Model of Aggregate Mental Score";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for AGG\_MENT

\*-----------------------------------------------------------;

\*First run: 2500 nbi, 30000 nmc;

\*DIC: 3330.507;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betaAM **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAM betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAM\*AGG\_MENT + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_aggment ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Mental Score";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for AGG\_MENT (remove drink)

\*-----------------------------------------------------------;

\*First run: 2500 nbi, 30000 nmc, DIC = 3328.351;

\*DIC improved w/o drink, remove it in next step;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betaAM **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAM betadrugs betaage betabmi betaadh betarace

betasmoke betamarij betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAM\*AGG\_MENT + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_aggment ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Mental Score (Remove Drink)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for AGG\_MENT (remove smoke)

\*-----------------------------------------------------------;

\*DIC: 3331.001, DIC slightly better w/ smoke, can keep;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betaAM **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAM betadrugs betaage betabmi betaadh betarace

betadrink betamarij betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAM\*AGG\_MENT + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_aggment ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Mental Score (Remove Smoke)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for AGG\_MENT (remove marij)

\*-----------------------------------------------------------;

\*First run: 2500 nbi, 30000 nmc, DIC: 3330.325;

\*DIC slightly worse w/ marij, remove in next step for parsimony;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betaAM **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAM betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAM\*AGG\_MENT + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_aggment ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Mental Score (Remove Marijuana)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for AGG\_MENT (remove income)

\*-----------------------------------------------------------;

\*DIC: 3449.451, much worse w/o income, keep it;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betaAM **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0** betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAM betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAM\*AGG\_MENT + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY + betaeduc\*educHSmore;

model diff\_aggment ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Mental Score (Remove Income)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for AGG\_MENT (remove educ)

\*-----------------------------------------------------------;

\*DIC: 3329.512, better w/o education, try to remove in next step;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betaAM **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**;

PARMS sigma2 **1**;

PRIOR betaAM betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaincmed betainchigh ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAM\*AGG\_MENT + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh;

model diff\_aggment ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Mental Score (Remove Education)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Next step: try to remove drink, marijuana, education

\*-----------------------------------------------------------;

\*DIC: 3327.321, better w/o these vars, can remove them;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betaAM **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betasmoke **0** betaincmed **0** betainchigh **0**;

PARMS sigma2 **1**;

PRIOR betaAM betadrugs betaage betabmi betaadh betarace

betasmoke betaincmed betainchigh ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAM\*AGG\_MENT + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betasmoke\*smokecurrent + betaincmed\*incomemed + betainchigh\*incomehigh;

model diff\_aggment ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Mental Score (Remove 3 vars)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Next step: see if smoke should be removed or not

\*-----------------------------------------------------------;

\*DIC: 3327.481;

\*Approximately the same, can remove for parsimony;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betaAM **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betaincmed **0** betainchigh **0**;

PARMS sigma2 **1**;

PRIOR betaAM betadrugs betaage betabmi betaadh betarace

betaincmed betainchigh ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAM\*AGG\_MENT + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betaincmed\*incomemed + betainchigh\*incomehigh;

model diff\_aggment ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Mental Score (Remove 3 vars + Smoke)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Crude model for agg\_phys (drugs + base only)

\*-----------------------------------------------------------;

\*DIC: 3320.175, DIC is better in full model :);

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC seed = **204**;

PARMS betaint **0** betaAP **0** betadrugs **0**;

PARMS sigma2 **1**;

PRIOR betaAP betadrugs ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAP\*AGG\_PHYS + betadrugs\*harddrugsY;

model diff\_aggphys ~ normal(mu, var = sigma2);

title "Crude Model of Aggregate Physical Score";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for agg\_phys

\*-----------------------------------------------------------;

\*First run: 2500 burning, 30000 its, DIC: 3094.218;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC seed = **204**;

PARMS betaint **0** betaAP **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAP betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAP\*AGG\_PHYS + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_aggphys ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Physical Score";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for agg\_phys (remove drink)

\*-----------------------------------------------------------;

\*First run: 2500 burning, 30000 its, DIC: 3092.302;

\*DIC improved w/o drink, remove it in next step;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC seed = **204**;

PARMS betaint **0** betaAP **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAP betadrugs betaage betabmi betaadh betarace

betasmoke betamarij betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAP\*AGG\_PHYS + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_aggphys ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Physical Score (Remove Drink)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for agg\_phys (remove smoke)

\*-----------------------------------------------------------;

\*DIC: 3092.216, DIC improved w/o smoke, remove in next step;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC seed = **204**;

PARMS betaint **0** betaAP **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAP betadrugs betaage betabmi betaadh betarace

betadrink betamarij betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAP\*AGG\_PHYS + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_aggphys ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Physical Score (Remove Smoke)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for agg\_phys (remove marijuana)

\*-----------------------------------------------------------;

\*DIC: 3092.388, DIC slightly better w/o marijuana, remove in next step;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC seed = **204**;

PARMS betaint **0** betaAP **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAP betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAP\*AGG\_PHYS + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_aggphys ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Physical Score (Remove Marijuana)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for agg\_phys (remove income)

\*-----------------------------------------------------------;

\*DIC: 3220.329, way worse w/o income, need to keep it;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC seed = **204**;

PARMS betaint **0** betaAP **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaAP betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAP\*AGG\_PHYS + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY +

betaeduc\*educHSmore;

model diff\_aggphys ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Physical Score (Remove Income)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for agg\_phys (remove education)

\*-----------------------------------------------------------;

\*DIC: 3093.707, slightly lower w/o education, remove it;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC seed = **204**;

PARMS betaint **0** betaAP **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**;

PARMS sigma2 **1**;

PRIOR betaAP betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaincmed betainchigh ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAP\*AGG\_PHYS + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh;

model diff\_aggphys ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Physical Score (Remove Education)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Model for agg\_phys (remove drink, smoke, marijuana, educ)

\*-----------------------------------------------------------;

\*DIC: 3089.996, dic is better now, remove all 4 of these;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC seed = **204**;

PARMS betaint **0** betaAP **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betaincmed **0** betainchigh **0**;

PARMS sigma2 **1**;

PRIOR betaAP betadrugs betaage betabmi betaadh betarace

betaincmed betainchigh ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaAP\*AGG\_PHYS + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betaincmed\*incomemed + betainchigh\*incomehigh;

model diff\_aggphys ~ normal(mu, var = sigma2);

title "Full Model of Aggregate Physical Score (Remove 4 vars)";

**RUN**; title;

\*All vars must stay in model now, keep it here;

\*-----------------------------------------------------------;

\* Crude model of just drugs for logvload (and baseline)

\*-----------------------------------------------------------;

\*DIC: 390.449 (better for full model);

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **50000** plots = all DIC;

PARMS betaint **0** betalogvload **0** betadrugs **0**;

PARMS sigma2 **1**;

PRIOR betalogvload betadrugs ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betalogvload\*logvload + betadrugs\*harddrugsY;

model diff\_logvload ~ normal(mu, var = sigma2);

title "Crude Model of log10 VLOAD (drugs only)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model of log10vload

\*-----------------------------------------------------------;

\*DIC: 372.812;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **50000** plots = all DIC;

PARMS betaint **0** betalogvload **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betalogvload betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betalogvload\*logvload + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_logvload ~ normal(mu, var = sigma2);

title "Full Model of log10 VLOAD";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model of log10vload (remove drink)

\*-----------------------------------------------------------;

\*DIC: 370.643, is better w/o drink, so remove drink;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **30000** plots = all DIC;

PARMS betaint **0** betalogvload **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betalogvload betadrugs betaage betabmi betaadh betarace

betasmoke betamarij betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betalogvload\*logvload + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_logvload ~ normal(mu, var = sigma2);

title "Full Model of log10 VLOAD (remove drink)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model of log10vload (remove smoke)

\*-----------------------------------------------------------;

\*DIC: 370.716, better w/o smoke, so remove smoke in next step;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **50000** plots = all DIC;

PARMS betaint **0** betalogvload **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betalogvload betadrugs betaage betabmi betaadh betarace

betadrink betamarij betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betalogvload\*logvload + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_logvload ~ normal(mu, var = sigma2);

title "Full Model of log10 VLOAD (remove smoke)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model of log10vload (remove marij)

\*-----------------------------------------------------------;

\*DIC: 371.730, is better w/o marij, so remove from next step;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **50000** plots = all DIC;

PARMS betaint **0** betalogvload **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betalogvload betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betaincmed betainchigh betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betalogvload\*logvload + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_logvload ~ normal(mu, var = sigma2);

title "Full Model of log10 VLOAD (remove marij)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model of log10vload (remove income)

\*-----------------------------------------------------------;

\*DIC: 384.030, worse w/o income, so keep income;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **50000** plots = all DIC;

PARMS betaint **0** betalogvload **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betalogvload betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaeduc ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betalogvload\*logvload + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY +

betaeduc\*educHSmore;

model diff\_logvload ~ normal(mu, var = sigma2);

title "Full Model of log10 VLOAD (remove income)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model of log10vload (remove educ)

\*-----------------------------------------------------------;

\*DIC: 370.557, better w/o educ, can remove in next step;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **50000** plots = all DIC;

PARMS betaint **0** betalogvload **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**;

PARMS sigma2 **1**;

PRIOR betalogvload betadrugs betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaincmed betainchigh ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betalogvload\*logvload + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh;

model diff\_logvload ~ normal(mu, var = sigma2);

title "Full Model of log10 VLOAD (remove educ)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Next model: remove drink, smoke, marij, educ in 1 step

\*-----------------------------------------------------------;

\*DIC: 368.099;

**PROC** **MCMC** data = hiv nbi = **2500** nmc = **50000** plots = all DIC;

PARMS betaint **0** betalogvload **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betaincmed **0** betainchigh **0**;

PARMS sigma2 **1**;

PRIOR betalogvload betadrugs betaage betabmi betaadh betarace

betaincmed betainchigh ~ normal(mean = **0**, var = **1000**);

PRIOR betaint ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betalogvload\*logvload + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betaincmed\*incomemed + betainchigh\*incomehigh;

model diff\_logvload ~ normal(mu, var = sigma2);

title "Full Model of log10 VLOAD (rem. 4 vars)";

**RUN**; title;

\*DIC is best w/o them, remove all 4 vars;

\*-----------------------------------------------------------;

\* Crude model for LEU3N for drugs only (and base)

\*-----------------------------------------------------------;

\*DIC: 6285.983, DIC is better w/ other things in model;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Crude Model of CD4+ count (drugs)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for LEU3N

\*-----------------------------------------------------------;

\*DIC: 5874.293;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaincmed betainchigh betaeduc betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for LEU3N (remove drink)

\*-----------------------------------------------------------;

\*DIC: 5878.247, lower w/o drink so remove at next step;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betasmoke **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betasmoke betamarij betaincmed betainchigh betaeduc betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betasmoke\*smokecurrent + betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (Remove drink)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for LEU3N (remove smoke)

\*-----------------------------------------------------------;

\*DIC: 5878.376, lower w/o smoke so can remove in next step;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betadrink betamarij betaincmed betainchigh betaeduc betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (Remove smoke)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for LEU3N (remove marij)

\*-----------------------------------------------------------;

\*DIC: 5882.963, is worse w/o marijuana, so keep it in at next step;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betadrink betasmoke betaincmed betainchigh betaeduc betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (Remove marijuana)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for LEU3N (remove income)

\*-----------------------------------------------------------;

\*DIC: 6122.414, definitely need to keep income;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaeduc betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY +

betaeduc\*educHSmore;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (Remove income)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Full model for LEU3N (remove educ)

\*-----------------------------------------------------------;

\*DIC: 6122.414, need to keep education at next step;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betadrink **0** betasmoke **0** betamarij betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betadrink betasmoke betamarij betaeduc betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW + betadrink\*drink13plus +

betasmoke\*smokecurrent + betamarij\*marijY +

betaeduc\*educHSmore;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (Remove educ)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Next model: remove drink and smoke together

\*-----------------------------------------------------------;

\*DIC: 5877.083 (better w/o drink/smoke, okay to remove both);

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betamarij **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betamarij betaincmed betainchigh betaeduc betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (rem. drink/smoke)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Next model: remove drink and smoke together

\* See if income can be removed

\*-----------------------------------------------------------;

\*Autocorrelation is bad w/o income and DIC increases a lot, keep income in;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betamarij **0** betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betamarij betaeduc betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betamarij\*marijY + betaeduc\*educHSmore;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (rem. drink/smoke + inc.)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Next model: remove drink and smoke together (and educ)

\*-----------------------------------------------------------;

\*DIC: 5877.100, v. slightly above, take out of model for parsimony;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betamarij **0** betaincmed **0** betainchigh **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betamarij betaincmed betainchigh betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betamarij\*marijY + betaincmed\*incomemed + betainchigh\*incomehigh;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (rem. drink/smoke + educ.)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Next model: remove drink and smoke together (and marij)

\*-----------------------------------------------------------;

\*DIC: 5880.650, can remove marijuana;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betaincmed **0** betainchigh **0**

betaeduc **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betaincmed betainchigh betaeduc betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betaincmed\*incomemed + betainchigh\*incomehigh +

betaeduc\*educHSmore;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (rem. drink/smoke + marij)";

**RUN**; title;

\*-----------------------------------------------------------;

\* Final model: remove 4 vars

\*-----------------------------------------------------------;

\*DIC: 5880.650, can remove marijuana;

**PROC** **MCMC** data = hiv nbi = **7500** nmc = **250000** thin = **10** plots = all DIC seed = **204**;

PARMS betaint **0** betaLEU3N **0** betadrugs **0** betaage **0** betabmi **0** betaadh **0**

betarace **0** betaincmed **0** betainchigh **0**;

PARMS sigma2 **1**;

PRIOR betaLEU3N betaage betabmi betaadh betarace

betaincmed betainchigh betaint betadrugs ~ normal(mean = **0**, var = **10000**);

PRIOR sigma2 ~ igamma(shape = **2.001**, scale = **1.001**);

mu = betaint + betaLEU3N\*LEU3N + betadrugs\*harddrugsY + betaage\*age +

betabmi\*BMI + betaadh\*adhhigh + betarace\*raceNHW +

betaincmed\*incomemed + betainchigh\*incomehigh;

model diff\_LEU3N ~ normal(mu, var = sigma2);

title "Full Model of CD4+ count (rem. 4 vars)";

**RUN**; title;

\*final model: remove drink, smoke, educ, marij;