**HAART Treatment Response in Hard Drug vs. Non Hard Drug Users**

**INTRODUCTION**

The Multicenter AIDS Cohort Study is interested in learning how effective highly active antiretroviral treatment (HAART) is between men who do and do not report using hard drugs at initiation of treatment. Specifically, the study is interested in treatment response between these two groups two years after initiating treatment. Prior animal HIV studies indicate that hard drug use limits the immune system response, however, the response in humans is unknown. While this study only observes treatment response from baseline to year two—it’s imperative to not individual response between those who use hard drugs at baseline and those that do not as a more aggressive treatment may be structured for those who are not responding for the future. The study observes overall if the subjects health improves by four measures, two of which are quality of life and two which are laboratory. In theory, viral load and T-cell counts have an opposing relationship, with viral load in recovering individuals declining. The quality of life measures are somewhat unknown as they should improve as health improves, but the aggressive nature of the treatment may have an adverse effect. Overall the hypotheses for this investigation include:

1. There is a difference in the effectiveness of HAART between hard drug and non-hard drug users to improve aggregate mental score at year two; the null is indicative of no difference between groups.
2. There is a difference in the effectiveness of HAART between hard drug and non-hard drug users to improve aggregate physical score at year two; the null is indicative of no difference between groups.
3. There is a difference in the effectiveness of HAART between hard drug and non-hard drug users to decrease viral load at year two; the null is indicative of no difference between groups.
4. There is a difference in the effectiveness of HAART between hard drug and non-hard drug users to increase T-cell count at year two; the null is indicative of no difference between groups.

Effectiveness of HAART is measured by two different time points: each outcome at baseline, and each outcome at year two. Adherence to HAART is also measured for efficacy of the study. This investigation was a retrospective cohort study with participants self-reporting the use of hard-drugs at baseline, with up to eight years of data on each participant.

**METHODS**

While the original data included 34 variables, variables for all outcomes were created to model the difference between baseline and year two. In addition, covariates were recoded for specific group interest. Covariates for each outcome model included: the outcome at baseline, age at baseline, BMI at baseline, race (non-white Hispanic or other), baseline marijuana use, baseline alcohol use (greater than 13 drinks per day or less than 13 drinks per day), baseline smoking status (current smoker or not), baseline salary (less than $10000 annually, $10000-$40000 annually, or greater than $40000 annually), education (some high school or high school completion and beyond), as well as HAART adherence (greater than 95% adherence or less than 95% adherence to treatment).

Outcome variables included two laboratory measures: viral load, and CD4+ T cell count, along with two quality of life measures: aggregate physical score, and aggregate mental score. The outcome, viral load was log transformed for clinical interpretation.

Observations without measurements for any of the outcome variables at baseline or year two were deleted, along with observations without a response for hard drug use. Additionally, extraneous BMI measures were deleted (either missing or negative measures). The final dataset included 397 observations, and 23 variables for analysis.

Data points for all outcomes at baseline and year two were plotted to ensure there were not any violations of independence, linearity, homoscedasticity, little multicollinearity, or multivariate normality. The final dataset was observed for missing data patterns and only resulted in one, with sixteen observations missing responses for salary; this was ignored in analysis as the missing data spread was similar across drug use groups.

Descriptive statistics were gathered on the data to include the mean, standard deviation and frequency for continuous variables; and the frequency and percent for categorical variables (regardless of their involvement in the final model).

Independent associations for each covariate against each outcome variable were run to determine significance. It’s important to note that the measure of each outcome variable at baseline were included as covariates in independent associations and the final models; as that those starting at significant losses in T-Cell count, aggregate physical score, aggregate mental score, or significant highs in viral load may see more of an effect than those within “normal” range values.

For model-building, four multivariable regression models were built using step-wise regression fitting. The full model for each outcome included: drug use, age, BMI, baseline measure of the outcome, baseline alcohol use, smoking status, race, salary, education, baseline marijuana use, and HAART adherence. Only variables that significantly contributed to the final model were kept (p<0.05). The final model for aggregate mental score included hard drug use group, aggregate mental score at baseline and race. The final model for aggregate physical score included hard drug use group, age, and aggregate physical score at baseline. The final model for T-cell count included hard drug use group, BMI, salary, and education. The final model for viral load included hard drug use group, viral load at baseline, race, salary, education, and HAART adeherence.

All programming was done using SAS 9.4 for Windows.

**RESULTS**

Between baseline and year two, treatment response to HAART was observed for 397 patients. Most patients were non-Hispanic whites (Table 2). Most patients made a salary between $10,000 and $40,000 annually (Table 2). Most patients were high-school educated (Table 2). Most patients were not current smokers, drank less than 13 drinks per week, had never used marijuana, and adhered to treatment greater than 95% of the time (Table 2). Mean baseline age between hard drug and non hard drug users was reportedly the same at 48.87 (5.68) and 45.17 (8.87) years respectively. Similarly with BMI, 24.65 (3.29) for hard drug users and 25.60 (4.12) for non hard drug users. Baseline scores for aggregate mental score, aggregate physical score, and viral load were reportedly similar between the two groups (Table 1). CD4+T-cell count, however, was significantly different: with hard drug users reporting a high count of 455.36, and non hard drug users reporting lower counts at 364.62 (Table 1).

Full results from the descriptive analysis can be referenced from Tables 1 and 2 in the Appendix.

The final model for aggregate mental score difference from baseline to year two showed a significant association between aggregate mental score at baseline (p <.0001) as well as race (p = 0.0033) to the outcome; but while hard drug use was significantly associated with the outcome, there was not a significant difference between drug use groups. The final model for aggregate physical score difference from baseline to year two showed a significant association between hard drug use (p = 0.0016), age (p = 0.0347), aggregate physical score at baseline (p <.0001), education (p = 0.0038) and HAART adherence (p = 0.0361). The final model for CD4+T-Cell count difference from baseline to year two showed a significant association between BMI (p = 0.0378), salary (p = 0.0654), and education (p = 0.0222). Although, hard drug use overall was significant, there was not a significant difference between treatment groups. The final model for viral load difference from baseline to year two showed a significant association for viral load at baseline (p <.0001), salary (p = 0.0225), education (p = 0.0007), and HAART adherence (p = 0.0083). While hard drug use was significant overall, only those who used hard drugs showed a significant association with the outcome.

Model fit for each outcome as compared from the full to final models was not significantly different, and thus the reduced models for each outcome were chosen to represent the final relationships.

It is interesting to note that the difference in physical score, and CD4+T-Cell count was higher among those who used drugs at baseline. While aggregate mental score difference and viral load difference was relatively the same between the two groups (Figure 1).

Full results from the multivariable regression analyses can be referenced from Table 3 in the Appendix.

**CONCLUSIONS**

It is important to note some limitations of these analyses include the fact that the drug use groups were significantly uneven. Additionally, loss to follow up overall reduced the sample size by a significant proportion of what it could have been, and it may be worth looking at the data as a whole (through year eight) to see if any other trends occur, or models could be adjusted for missing year measures. Due to aggregate mental score and aggregate physical score being a combination of factors, it seems logical as to why there was no difference between drug use groups with aggregate mental score, although there was with aggregate physical score. This could explain the presumed relationship that treatment is mentally taxing on patients—and thus, contributes to a greater effect than drug use. The aggregate physical score, though, being significantly associated and different between drug use groups marks a relationship between the effectiveness of HAART on the body couple with the negative effects from hard drug use. Finally, the aggregate physical score outcome coupled with the viral load outcome and their observed differences between drug use groups should have investigators consider more aggressive treatment for those that use hard drugs. Seeing as how the overall mental quality is not particularly effected—it could be useful and important to consider more aggressive treatments in mean who suffer from this disease but don’t see as significant results as those who do not use hard drugs. Whether more aggressive treatment is a significant factor, or observing more time points from treatment overall, or working to reduce repeated drug use among these patients: an alternative should be reviewed for the overall health and success of treatment for those participating in the study.

**REPRODUCIBLE RESEARCH**

/\*Importing the data\*/

**PROC** **IMPORT** OUT= WORK.hiv

DATAFILE= "C:\Users\Kayla\Desktop\hiv\_6623\_final.csv"

DBMS=CSV REPLACE;

GETNAMES=YES;

DATAROW=**2**;

**RUN**;

The code for the final models can be seen here with additional code on GITHUB:

**PROC** **GLM** DATA = WORK.HIVCLEAN; /\*Final model for DMENTAL\*/

CLASS NHW;

MODEL DMENTAL = MENTALBASE NHW / SOLUTIONS;

**RUN**;

**PROC** **GLM** DATA = WORK.HIVCLEAN; /\*Final model for DPHYSICAL\*/

CLASS HARD\_DRUGS EDUC HARTADH;

MODEL DPHYSICAL = AGE HARD\_DRUGS PHYSICALBASE EDUC HARTADH / CLPARM SOLUTIONS;

**RUN**;

**PROC** **GLM** DATA = WORK.HIVCLEAN; /\*Final model for DLEU\*/

CLASS HARD\_DRUGS SALARY EDUC;

MODEL DLEU = HARD\_DRUGS BMINUM SALARY EDUC / CLPARM SOLUTIONS;

**RUN**;

**PROC** **GLM** DATA = WORK.HIVCLEAN; /\*Final model for DVLOAD\*/

CLASS HARD\_DRUGS NHW SALARY EDUC HARTADH BASEMJUSE;

MODEL DVLOAD = HARD\_DRUGS VLOADBASE NHW SALARY EDUC HARTADH / CLPARM SOLUTIONS;

**RUN**;

A link to the GITHUB repository where you can view full code can be found here: https://github.com/BIOS6623-UCD/bios6623-kbell28k/tree/master/Project1/Code

**APPENDIX**





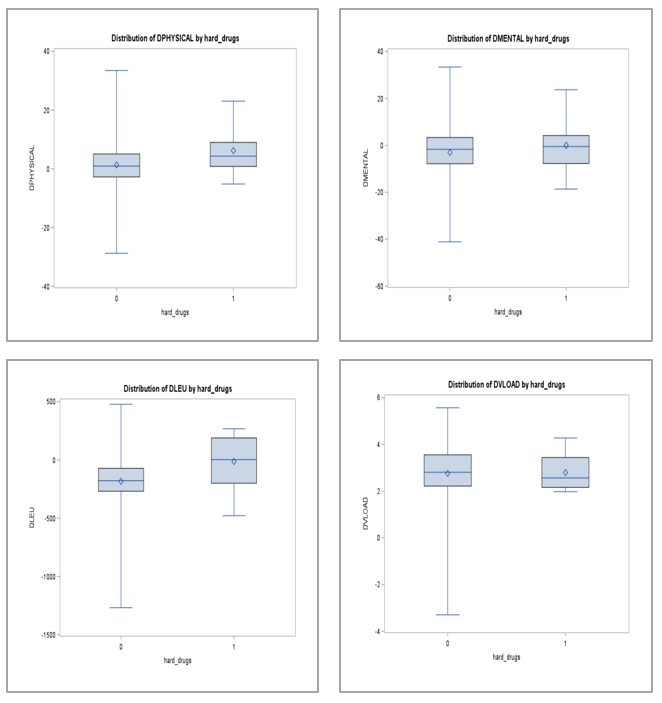


Figure 1. Distributions of difference in outcome variables by drug use.

