Project 1 Analysis Plan

Joseph Froelicher

September 27, 2021

Introduction

This is a secondary data analysis of the Multicenter AIDS Cohort Study, which an ongoing prospective cohort study of the natural and treated histories of HIV-1 infection in homosexual and bisexual men, in four major cities in the United States. Highly active anti-retroviral treatment (HAART) is the standard treatment for HIV infected patients. The data include up to 8 years of longitudinal laboratory and quality of life measures, in addition to demographic and other health information, on HIV infected men after beginning HAART. The subjects were seen annually. Year 0 data are from the subjects' last untreated visit, just before beginning HAART. All other visits (year 1 up to 8) are on treatment.

There are four measures of treatment response. The first two are laboratory measures, viral load (VLOAD), which is the number of HIV copies in one mL of blood, and the second, is CD4+ T cell count (LEU3N), a measure of immunologic health. In untreated HIV infection, viral load increases over time and CD4+ T cell counts decline as the immune system is attacked by the virus. Once treatment is initiated, the investigators expect viral load to decrease rapidly and CD4 counts to recover. The final two measures are quality of life measures from the SF-36. The first is the aggregate physical quality of life score (AGG_PHYS) and the second is the aggregate mental quality of life score (AGG_MENT). These scores range from 0 to 100, with higher scores indicating better quality of life. It is unclear what happens to quality of life after initiating treatment. Theoretically, subjects' improving health should result in increased quality of life, however, the side effects of these treatments are significant. If subjects experience declines in quality of life after initiating treatment there should be concern that subjects' would stop treatment.

The purpose of this report is to develop the analysis plan to answer the following hypothesis/question of interesd:

1. How is treatment response 2 years after initiating HAART differerent between subjects who report using hard drugs, such as heroine and cocaine, at baseline and other subjects, who did not report hard drug use at baseline.
2. Are any additional variables are associated with hard drug use. The investigators have identified several variables of interest, including: age, BMI, smoking status, education, race/ethnicity, and adherence.

Methods

Data Cleaning

Value ranges have been provided for numeric variables, and have been replaced as missing. Viral load will be log base 10 transformed due to extremely large values. Typical ranges of Cd4+ are between 500 and 1,500, and thus a value below 200 results in an AIDS Diagnosis. This will be kept in mind during the analysis. Values of 9 or 999 were recorded for DYSLIP, FP, FRP, BMI, where there were insufficient data, and thus will be treated as missing.

Data Analysis

The analysis performed will be a simple linear regression in a Bayesian framework, with its frequentist counterpart used as a benchmark for analysis. The analysis will examine the four quality of life measures at baseline and 2 years. Each of the variables of interest, Age, BMI, smoking status, education, and race/ethnicity are recorded at baseline. The variable of interest, adherence, is recorded at 2 years. The analyst intends to do backwards selection to identify any associations between quality of life and the variables of interest.

Preliminary Results

	Hard Drugs	Absence	Presence
n		3611	291
Mental Score (Mean ± sd)		47.32 ± 12.54	43.34 ± 14.26
Physical Score (Mean ± sd)		50.12 ± 9.79	46.52 ± 10.67
BMI (Mean ± sd)		36.26 ± 101.25	67.89 ± 201.62
Log Viral Load (Mean ± sd)		2.21 ± 1.60	2.29 ± 1.71
CD4 positive cell count (Mean ± sd)		546.14 ± 276.77	448.17 ± 268.19
Age (Mean ± sd)		46.47 ± 9.16	46.65 ± 8.85
Adherence (%)	Not Adherent	2633 (88.9)	177 (78.7)
	Adherent	330 (11.1)	48 (21.3)
Race/Ethnicity (%)	White, non-Hispanic	2357 (65.3)	142 (48.8)
	White, Hispanic	146 (4.0)	9 (3.1)
	Black, non-Hispanic	889 (24.6)	115 (39.5)
	American Indian/Alaska Native	31 (0.9)	6 (2.1)
	Asian or Pacific Islander	32 (0.9)	1 (0.3)
	Other	156 (4.3)	18 (6.2)
Education (%)	8th grade or less	32 (0.9)	0 (0.0)
	9, 10, or 11th grade	156 (4.3)	46 (15.8)
	12th grade	535 (14.8)	53 (18.2)
	some college, no degree	1305 (36.1)	100 (34.4)
	Bachelor's degree	810 (22.4)	19 (6.5)
	some graduate school	258 (7.1)	6 (2.1)
	Post-graduate degree	515 (14.3)	67 (23.0)
Smoker (%)	Never	949 (26.3)	52 (17.9)
	Former	1457 (40.3)	55 (18.9)
	Current	1205 (33.4)	184 (63.2)