BIOS 6623: Project 2

Fall 2024

Written report, source codes and in-class presentation due date:

1PM on Monday, October 21

Data Overview: This project is a secondary data analysis of the Multicenter AIDS Cohort Study, an ongoing prospective cohort study of the natural and treated histories of HIV-1 infection in homosexual and bisexual men in 4 major cities in the United States.

Highly active antiretroviral therapy (HAART) is the standard treatment for HIV infected patients. Our dataset includes up to 8 years of longitudinal laboratory and quality of life measures, as well as 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.

Question of Interest: We are interested in understanding how treatment response 2 years after initiating HAART differs between subjects who report using hard drugs, such as heroin and cocaine, and other subjects, who did not report hard drug use. We would like to compare subjects who never report using hard drugs to subjects that report using hard drugs at year 2 (current hard drug users), as well as to subjects that did not use hard drugs at the year 2 visit but reported hard drug use in the past (previous hard drug use, reported at either or both of the year 0 and 1 visits). **We would also like to understand if differences in treatment response between the drug use groups can be explained by differences in adherence to the HAART treatment regimen.**

Biologic Motivation: There is limited evidence from laboratory in vitro and animal studies that the use of hard drugs inhibits the immune system and increases HIV replication; however, results have not been clear in human studies. If drug users have poor treatment response compared to others, we may need to consider more aggressive treatment strategies or more actively encourage patients to enroll in drug rehabilitation programs.

Information about variables: We have 4 measures of treatment response. The first two are laboratory measures, viral load (VLOAD), which is the number of HIV copies in a 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, we expect viral load to decrease rapidly and CD4 counts to recover. Our last 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. We are not sure what happens to quality of life after initiating treatment. While in theory subjects' improving health should result in increased quality of life, the side effects of these treatments are significant. If subjects experience declines in quality of life after initiating treatment, we would be concerned that they would stop treatment.

The data file of Project 2 is **hiv_dataset.csv**, which can be access on Canvas.
A data dictionary can also be found on Canvas: **codebook_6623_hiv.pdf**.