libname vacsdata 'file path';

proc import datafile='/file path/vacs\_scd\_dat.csv'

out=vacs dbms=csv replace;

guessingrows=MAX;

run;

data vacs;

set vacs;

format baseline date\_death end\_followup mmddyy10.;

baseline = input(put(baseline, $10.), mmddyy10.);

date\_death = input(put(date\_death, $10.), mmddyy10.);

end\_date = '31DEC2014'd;

if missing(date\_death) then end\_followup = end\_date;

else if date\_death < end\_date then end\_followup = date\_death;

else end\_followup = end\_date;

followup\_days = end\_followup - baseline;

followup\_years = followup\_days / 365.25;

run;

data vacs;

set vacs;

length CD4\_category VL\_category 8;

if hiv = 0 then CD4\_category = 0;

else if cd4 >= 500 then CD4\_category = 1;

else if 200 <= cd4 < 500 then CD4\_category = 2;

else if cd4 < 200 then CD4\_category = 3;

if hiv = 0 then VL\_category = 0;

else if vl < 500 then VL\_category = 1;

else if vl >= 500 then VL\_category = 2;

run;

data vacs;

set vacs;

Black = (racecomg = 2);

Hispanic = (racecomg = 3);

Other = (racecomg = 4);

run;

proc means data=vacs n nmiss;

run;

proc freq data=vacs;

tables hiv;

run;

proc means data=vacs n mean std median q1 q3;

class hiv;

var age egfr bmi hgb;

run;

%macro cat\_summary(var);

proc freq data=vacs;

tables hiv\*&var / chisq;

run;

%mend;

%cat\_summary(female);

%cat\_summary(htn);

%cat\_summary(CD4\_category);

%cat\_summary(VL\_category);

/\* PART 2: Kaplan-Meier Survival and Cox Models \*/

proc lifetest data=vacs plots=survival(atrisk=0 to 10 by 2);

time followup\_years \* inc\_scd(0);

strata hiv;

run;

proc phreg data=vacs;

class hiv(ref='0');

model followup\_years \* inc\_scd(0) = hiv;

run;

/\* PART 3: Poisson Regression \*/

proc genmod data=vacs;

class hiv(ref='0') Black Hispanic Other / param=ref;

model inc\_scd = hiv age female Black Hispanic Other / dist=poisson link=log offset=log\_followup;

/\* Need to create log\_followup first \*/

run;

/\* Create log(followup\_years) for offset \*/

data vacs;

set vacs;

log\_followup = log(followup\_years);

run;

/\* PART 4: Multiple Imputation (requires MI + MIANALYZE) \*/

proc mi data=vacs nimpute=5 out=mi\_data seed=82508;

var htn dyslipidemia smk\_stat egfr bmi hgb hcv3 CD4\_category VL\_category followup\_years age female hiv Black Hispanic Other prev\_cvd dmcom alcohol cocaine copd inc\_scd regimen;

run;

proc phreg data=mi\_data;

class hiv(ref='0') Black Hispanic Other / param=ref;

model followup\_years \* inc\_scd(0) = hiv age female Black Hispanic Other;

by \_Imputation\_;

ods output ParameterEstimates=estimates;

run;

proc mianalyze parms=estimates;

modeleffects hiv age female Black Hispanic Other;

run;

/\* PART 5: Time-Varying Cox Regression \*/

proc import datafile='/file path/timevar\_cd4\_final.csv'

out=timevar dbms=csv replace;

guessingrows=MAX;

run;

data timevar;

set timevar;

if hiv = 0 then CD4\_category = 0;

else if cd4\_count >= 500 then CD4\_category = 1;

else if 200 <= cd4\_count < 500 then CD4\_category = 2;

else if cd4\_count < 200 then CD4\_category = 3;

Black = (racecomg = 2);

Hispanic = (racecomg = 3);

Other = (racecomg = 4);

run;

proc phreg data=timevar;

class CD4\_category(ref='0') Black Hispanic Other / param=ref;

model (tstart, tstop) \* scd(0) = CD4\_category age female Black Hispanic Other alcohol cocaine htn dmcom dyslipidemia egfr bmi hgb prev\_cvd hcv3;

run;

/\* PART 6: Recurrent Event Model \*/

proc import datafile='/file path/recurrent\_mi\_vacs.csv'

out=recurrent dbms=csv replace;

guessingrows=MAX;

run;

data recurrent\_clean;

set recurrent;

if tstop > tstart;

if racecomg = 2 then Black = 1; else Black = 0;

if racecomg = 3 then Hispanic = 1; else Hispanic = 0;

if racecomg = 4 then Other = 1; else Other = 0;

run;

proc phreg data=recurrent\_clean;

class hiv(ref='0') Black Hispanic Other / param=ref;

model (tstart, tstop) \* ami(0) = hiv age\_bl female Black Hispanic Other;

run;