libname smoke "C:\Users\petez\Desktop\SAS";

**run**;

libname p2 'C:\users\eprice\Desktop\576C\P2';

**proc** **contents** data=p2.frmgham;

**run**;

/\* risk factors:

sex period time age sysbp diabp bpmeds cursmoke cigpday totchol hdlc ldlc

bmi glucose diabetes heartrte prevap prevchd prevmi prevstrk prevhyp

\*/

**proc** **freq** data = p2.frmgham;

table angina\*period hospmi\*period mi\_fchd\*period anychd\*period stroke\*period cvd\*period hyperten\*period / norow nocol nopercent;

**run**;

**proc** **freq** data = p2.frame;

table period\*prevap period\*prevmi / nocol norow nopercent;

**run**;

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\*create create start & end times for each period exmtime & endtime (use code from study description)

create unique incidence event for MI and Angina\*/

**data** Uframe;

set tmp1.frmgham;

if prevap = **1** then delete;

if prevmi = **1** then delete;

**run**;

**proc** **sort** data=uframe;

by randid descending period;

**run**;

/\*create start & end time frames for each period exmtime & endtime\*/

**data** uframe;

set uframe;

by randid;

retain exmtime;

/\* identify final follow-up time\*/

final\_fu= max(timeap, timemi, timemifc, timechd, timestrk, timecvd, timehyp);

/\*create time frame for each period\*/

if first.randid then do;

endtime=final\_fu; exmtime=time;

end;

else do;

endtime=exmtime;

exmtime=time;

end;

**run**;

/\*if angina or MI eventtime happened between period exmtime and period endtime, unique\_ event =1\*/

/\* this is still weird-- not to include fatal chd, at least... but following prof's definition and only counting MI & angina\*/

**data** uframe;

set uframe;

unique\_mi = **0**;

unique\_ang= **0**;

bad\_heart=**0**;

\*if angina = 1 and hospmi = 1 then do;

\* first\_time = min(timeap,timemi);

if hospmi =**1** and timemi > exmtime and timemi <= endtime then unique\_mi=**1**;

if angina =**1** and timeap > exmtime and timeap <= endtime then unique\_ang=**1**;

if unique\_mi eq **1** or unique\_ang eq **1** then bad\_heart =**1**;

**run**;

**proc** **sort** data=uframe;

by randid period;

**run**;

**proc** **freq** data=frame1;

tables unique\_ang unique\_mi bad\_heart;

**run**;

/\*checks that there are no subjects that have angina and hospmi first events that happen in different periods\*/

**data** timeck;

set uframe;

if angina = **1** and hospmi = **1** and unique\_mi ne unique\_ang;

**run**;

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\* data cleaning for time period 1 \*/

**data** frame1;

set uframe;

/\*cleaning sex \*/

if Sex = **2** then sex = **0**;

/\* only looking at period 1 \*/

if period = **2** or period = **3** then delete;

/\*getting rid of all observations that already came in w/ angina or MID \*/

if prevap = **1** then delete;

if prevmi = **1** then delete;

if prevchd =**1** then delete;

\*keep bad\_heart sex age sysbp diabp bpmeds cursmoke cigpday totchol hdlc ldlc

bmi glucose diabetes heartrte prevchd prevstrk prevhyp endtime exmtime unique\_ang unique\_mi;

**run**;

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\* data cleaning for time period 2 \*/

**data** frame2;

set uframe;

/\*cleaning sex \*/

if Sex = **2** then sex = **0**;

/\* only looking at period 2 \*/

if period = **1** or period = **3** then delete;

/\*getting rid of all observations that are prevalent for angina or MID \*/

if prevap = **1** then delete;

if prevmi = **1** then delete;

if prevchd =**1** then delete;

\*keep bad\_heart sex age sysbp diabp bpmeds cursmoke cigpday totchol hdlc ldlc

bmi glucose diabetes heartrte prevchd prevstrk prevhyp endtime exmtime unique\_ang unique\_mi;

**run**;

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\* data cleaning for time period 3 \*/

**data** frame3;

set uframe;

/\*cleaning sex \*/

if Sex = **2** then sex = **0**;

/\* only looking at period 3 \*/

if period = **1** or period = **2** then delete;

/\*getting rid of all observations that are prevalent for angina or MID \*/

if prevap = **1** then delete;

if prevmi = **1** then delete;

if prevchd =**1** then delete;

\*keep bad\_heart sex age sysbp diabp bpmeds cursmoke cigpday totchol hdlc ldlc

bmi glucose diabetes heartrte prevchd prevstrk prevhyp endtime exmtime unique\_ang unique\_mi;

**run**;

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

**proc** **corr** data = frame3;

var sex age sysbp diabp bpmeds cursmoke cigpday totchol

bmi glucose diabetes heartrte prevstrk prevhyp;

**run**;

**proc** **logistic** descending data=frame1;

model bad\_heart = sex age sysbp diabp cigpday totchol

bmi diabetes heartrte prevstrk prevhyp;

**run**;

/\*missing data from: cigpday-29 totchol-50 bmi-19 glucose-388 hrtrate-1 bpmeds-59 \*/

**proc** **freq** data = frame1;

tables sex age sysbp diabp cigpday totchol

bmi diabetes heartrte prevstrk prevhyp / nocol norow nopercent;

**run**;

/\*\*\*\*\*\*\*\*/

/\*Does quitting smoking reduce cvd\*/

/\*identify prevalent smoking population. identify quitters among smoking population\*/

/\*sort by ascending period!\*/

**proc** **sort** data= uframe;

by randid period;

**run**;

**data** smoke;

set uframe;

by randid;

retain smoker1;

if first.randid then do;

if cursmoke eq **1** then do;

smoker1 =**1**;

quitter =**0**;

end;

else smoker1 = **0**;

end;

else do;

if smoker1 eq **1** then do;

if cursmoke eq **0** then do;

quitter = **1**;

smoker1 = **0**;

end;

else quitter = **0**;

end;

end;

**run**;

**data** smoke1;

set smoke;

if quitter ne **.**;

**run**;

**proc** **freq** data=smoke1 ;

tables quitter;

**run**;