

```

/*SAMPLE SAS CODE FOR AIDS ELECTRONIC SUPPLEMENT*/
options ps=60 ls=80 nodate pageno=1;
dm log "clear;" continue; dm out "clear;" continue;

*set directory;
libname wihs "H:\Survival Tutorial\WIHS";

/*The following describes the example data included:
newid: participant id (this has been changed from original WIHS study id)
t: time to event (AIDS/Death) in years
taidsdth: time to AIDS/Death
tarv: time to HAART initiation
ageatfda: age in years at baseline
BASEIDU: 0-no history of IDU; 1- history of IDU
black: 0-not African-American; 1- African-American
cd4nadir: CD4 count nadir prior to baseline
*/

*read and prep data;
proc import datafile="H:\Survival Tutorial\WIHS\WIHS_IDU.csv"
out=prea dbms=csv replace;
    getnames=yes;
    datarow=2;
run;

data a; set prea(drop=t);
    pid=_n_;
    if black=1 then aa=1; else if black=0 then aa=0;
    if dth=1 then y=1; else if dth=0 then y=0; *death/AIDS indicator;
    label y="AIDS/death=1";
    idu=baseidu;
    age=ageatfda/10; *decades of age at index;
    cd4n=cd4nadir/100; *centiles of CD4 at index;
    t=round(taidsdth,.0001);
    if t>10 then do; t=10; y=0; end; *admin censor at 10 years;
    if t<9 and y=0 then drop=1; else drop=0; *indicator for drop out;
    label aa="african-am" age="decades" cd4n="centiles"
        y="AIDS/death" t="years" drop="drop out";
    /*time varying ARV exposure*/
    /*if they initiated arv before aids/death/censor then arv=1*/
    if tarv<taidsdth then arv=1; else arv=0;
    tarv=round(tarv,.0001);
    keep pid aa age cd4n idu y t drop arv tarv;
run;

```

```

*View the dataset;
proc print data=a(obs=10) noobs; var pid aa age cd4n idu y t drop arv tarv;
    title1 "First 10 rows of data";
run;

*Table 1;
proc means data=a fw=7 maxdec=3 n median q1 q3 min max mean sum;
    var age cd4n t;
    title1 "Table 1, overall";
proc freq data=a; tables aa idu arv drop y /missing;
proc sort data=a; by idu;
proc means data=a fw=7 maxdec=3 n median q1 q3 min max mean sum;
    var age cd4n t; by idu;
    title1 "Table 1, by idu";
proc freq data=a; tables idu*aa idu*y idu*arv/missing;
run;

*how many distinct event times?;
proc sort data=a; by idu t;
data distinct; set a; by idu t; where y=1; if last.t then output;
proc means data=distinct n; var t; title1 "No. distinct event times";
run;

*crude Cox model and curves;
proc means data=a n sum; var y t; title1 "No. cases and PY";
proc means data=a n sum; var y t; title1 "No. cases and PY";
where idu=1;
proc means data=a n sum; var y t; title1 "No. cases and PY";
where idu=0;
run;

*overall estimate of 10 year risk (1-S);
proc phreg data=a; model t*y(0)=/ties=efron rl; title1 "Crude Cox model";
    baseline out=overall survival=s lower=lcl upper=ucl LOGLOGS=logch;
data overall2;
set overall;
by t;
risk=1-s;
riskl=1-ucl;
risku=1-lcl;
if t=9.9274 then do; t=10; output; end;
proc print data=overall2;
run;

*unadjusted model;
proc phreg data=a; model t*y(0)=idu/ties=efron rl; title1 "Unadjusted Cox model";
    ods select modelinfo parameterestimates fitstatistics;

```

```

proc phreg data=a; model t*y(0)=idu idut/ties=efron rl; idut=idu*t; title1 "Assess PH
assumption";
    ods select modelinfo parameterestimates fitstatistics;
proc phreg data=a noprint; model t*y(0)=/ties=efron rl; strata idu;
    baseline out=crude survival=s lower=lcl upper=ucl cumhaz=ch LOGLOGS=logch;
data crude; set crude; by idu; output; if last.idu then do; t=10; output; end;
proc print data=crude noobs; where t=10; var idu s;
    title1 "10-year unadjusted AIDS-free survival by idu";
    run;

/*unadjusted figures*/
proc format;
value yes1fmt 1='Yes'
    0='No';
data crude;
set crude;
label idu='ID use';
format idu yes1fmt.;
run;
goptions reset=all device=zpng ftext="Albany AMT" htext=12pt gsfname=grafout
gsfmode=replace xmax=4 ymax=4 xpixels=4000 ypixels=4000; *1000dpi;
filename grafout "H:\Survival Tutorial\WIHS\unadjusted.png";
axis1 label=(angle=90 "Probability of AIDS free survival") order=(0 to 1 by .2) w=8 major=(w=8
h=15) minor=none offset=(0,0);
axis2 label=("Years") order=(0 to 10 by 2) w=8 major=(w=8 h=7) minor=none offset=(0,0);
symbol1 c=black v=none i=stepjs w=12 l=1;
symbol2 c=black v=none i=stepjs w=12 l=22;
Legend1 frame
    across=1 down=2
    position = (bottom left inside)
    offset = (10 pct,10 pct) shape=line(1 cm);
proc gplot data=crude; plot s*t=idu/vaxis=axis1 haxis=axis2 noframe nolegend; run; quit;

*log cumulative hazard plot;
goptions reset=all device=zpng ftext="Albany AMT" htext=12pt gsfname=grafout
gsfmode=replace xmax=4 ymax=4 xpixels=4000 ypixels=4000; *1000dpi;
filename grafout "H:\Survival Tutorial\WIHS\unadjusted_logch2.png";
axis1 label=(angle=90 "Log cumulative hazard") order=(-1.45 to 0.03 by .2) w=8 major=(w=8
h=15) minor=none offset=(0,0);
axis2 label=("Years") order=(0 to 10 by 2) w=8 major=(w=8 h=7) minor=none offset=(0,0);
symbol1 c=black v=none i=stepjs w=12 l=1;
symbol2 c=black v=none i=stepjs w=12 l=22;
Legend1 frame
    across=1 down=2

```

```

position = (top left inside)
offset = (2 pct,2 pct) shape=line(1 cm);
proc gplot data=crude; plot logch*t=idu/vaxis=axis1 haxis=axis2 noframe nolegend; run; quit;

```

```

*Get Unadjusted RD and RR;
data c1;
set crude;
if t=10 and idu=1;
data c2;
set crude;
if t=10 and idu=0;
data crude2;
merge c1(rename=(s=s_yes lcl=lcl_yes ucl=ucl_yes)) c2(rename=(s=s_no lcl=lcl_no ucl=ucl_no));
by t;
R_no=1-S_no;
Rnol=1-ucl_no;
Rnou=1-lcl_no;
R_yes=1-S_yes;
Ryesl=1-ucl_yes;
Ryesu=1-lcl_yes;
RD=R_yes-R_no;
drop idu t;
run;
proc print data=crude2;
run;

```

```

*IP weighted Cox model and curves;
*read spline macro code;
*Fit restricted quadratic splines on continuous variables;
%inc "H:\Survival Tutorial\WIHS\qspline4.sas";
%qspline4(a,age);
%qspline4(a,cd4n);

```

```

/*Model for exposure weights*/
/*Model Pr(IDU=1)*/
/*SINCE IDU IS AT BASELINE, WE DO NOT INCLUDE TIME UPDATED COVS*/
/*include all pairwise interactions*/
proc sort data=a; by pid;
proc logistic data=a noprint desc; model idu=; output out=n p=n;
proc logistic data=a desc; model idu=aa age age1-age3 cd4n cd4n1-cd4n3
aa*age aa*age1 aa*age2 aa*age3
aa*cd4n aa*cd4n1 aa*cd4n2 aa*cd4n3
age*cd4n age1*cd4n1 age1*cd4n2 age1*cd4n3

```

```

age2*cd4n2 age2*cd4n3
age3*cd4n3/lackfit; output out=d p=d;
ods select modelinfo parameterestimates fitstatistics lackfitchisq lackfitpartition;
run;
title1 "IP exposure model/denominator";
data a; merge a n d; by pid; if idu then ew=n/d; else ew=(1-n)/(1-d); label n= d=; drop _level_;
*Table showing imbalance by measured covariates fixed in weighted data;
proc logistic data=a desc; model idu=aa age cd4n/rl;
ods select modelinfo parameterestimates fitstatistics cloddswald;
title1 "Imbalance, observed";
proc logistic data=a desc; model idu=aa age cd4n/rl; weight ew;
ods select modelinfo parameterestimates fitstatistics cloddswald;
title1 "Imbalance, after weighting";
run;
*Aside: just controlling for confounding not selection;
proc phreg data=a covs; model t*y(0)=idu/ties=efron rl; weight ew;
title1 "IP (exposure only) weighted Cox model";
ods select modelinfo parameterestimates fitstatistics;
run;
*Drop out weights;
*Create dataset with monthly records for each subject (Hernan 2001);
data b; set a;
array j{121} j1-j121 (0,0.08,0.17,0.25, 0.33, 0.42, 0.5, 0.58, 0.67, 0.75, 0.83, 0.92,
1,1.08,1.17,1.25, 1.33, 1.42, 1.5, 1.58, 1.67, 1.75, 1.83, 1.92,
2,2.08,2.17,2.25, 2.33, 2.42, 2.5, 2.58, 2.67, 2.75, 2.83, 2.92,
3,3.08,3.17,3.25, 3.33, 3.42, 3.5, 3.58, 3.67, 3.75, 3.83, 3.92,
4,4.08,4.17,4.25, 4.33, 4.42, 4.5, 4.58, 4.67, 4.75, 4.83, 4.92,
5,5.08,5.17,5.25, 5.33, 5.42, 5.5, 5.58, 5.67, 5.75, 5.83, 5.92,
6,6.08,6.17,6.25, 6.33, 6.42, 6.5, 6.58, 6.67, 6.75, 6.83, 6.92,
7,7.08,7.17,7.25, 7.33, 7.42, 7.5, 7.58, 7.67, 7.75, 7.83, 7.92,
8,8.08,8.17,8.25, 8.33, 8.42, 8.5, 8.58, 8.67, 8.75, 8.83, 8.92,
9,9.08,9.17,9.25, 9.33, 9.42, 9.5, 9.58, 9.67, 9.75, 9.83, 9.92,
10);
do k=1 to 120;
in=j(k);
if j(k)<t<=j(k+1) then do; out=t; yy=y; drop2=drop; output; end;
else if j(k+1)<t then do; out=j(k+1); yy=0; drop2=0; output; end;
end;
keep pid in out yy idu aa age cd4n drop2 ew arv tarv;
run;

proc print data=b (obs=50);
var pid pid in out yy idu aa age cd4n drop2 ew arv tarv;
run;

```

```

/*Add restricted quadratic splines to continuous variables in dataset*/
%qspline4(b,age);
%qspline4(b,cd4n);
%qspline4(b,in);
%qspline4(b,out);
data b;
set b;
/*correct time varying covariate*/
/*so arv=1 only after initaition*/
if tarv>=out then arv=0;
    else if tarv<out then arv=1;
run;

*get person-time for IPW models;
*these are person months;
proc means data=b n sum; var yy out; title1 "IPW No. cases and PY";
proc means data=b n sum; var yy out; title1 "IPW No. cases and PY";
where idu=1;
proc means data=b n sum; var yy out; title1 "IPW No. cases and PY";
where idu=0;
run;

*Pr(Drop=0) is modeled;
*Time is modeled as continuous with a spline;
proc sort data=b; by pid in;
proc logistic data=b noprint; model drop2=idu in in1 in2 in3; output out=dn p=dn;
proc logistic data=b;
    model drop2=idu in in1 in2 in3
        aa age age1-age3 cd4n cd4n1-cd4n3 arv
        idu*aa idu*age idu*age1 idu*age2 idu*age3
            idu*cd4n idu*cd4n1 idu*cd4n2 idu*cd4n3 idu*arv
            aa*age aa*age1 aa*age2 aa*age3
            aa*cd4n aa*cd4n1 aa*cd4n2 aa*cd4n3
            age*cd4n age1*cd4n1 age1*cd4n2 age1*cd4n3
            age2*cd4n2 age2*cd4n3
            age3*cd4n3
            arv*aa arv*age arv*age1 arv*age2 arv*age3
            arv*cd4n arv*cd4n1 arv*cd4n2 arv*cd4n3 /lackfit; output out=dd p=dd;
ods select modelinfo parameterestimates fitstatistics lackfitchisq lackfitpartition;
title1 "IP drop-out model/denominator";
run;
data b; merge b dn dd; by pid in;
retain num den;
if first.pid then do; num=1; den=1; end;

```

```

        if in>0 then do;
            num=num*dn;
            den=den*dd;
        end;
        dw=num/den;
        fw=ew*dw;
        label dn= dd=; drop _level_;
        drop num den;
proc means data=b n mean std min max sum fw=8; var dn dd dw ew fw;
run;

*create table with sample records;
data b6;
set b;
in2=round(in,0.001);
out2=round(out,0.001);
stw2=round(ew,0.001);
censw2=round(dw,0.001);
finalw2=round(fw,0.001);
run;
ods rtf file="H:\Survival Tutorial\WIHS\patient_table.rtf";
proc print data=b6 ;
var pid in2 out2 idu yy drop2 stw2 censw2 finalw2;
where pid in (34,36,37,38,66);
run;
ods rtf close;

*Robust estimate of variance used with covs option;
proc phreg data=b covs; model out*yy(0)=idu/entry=in ties=efron rl;
    weight fw; title1 "IP weighted Cox model";
    ods select modelinfo parameterestimates fitstatistics;
run;
proc phreg data=b; model out*yy(0)=idu idut/entry=in ties=efron rl; weight fw;
    idut=idu*out; title1 "IPW model Assess PH assumption";
proc phreg data=b noprint ; model out*yy(0)=/entry=in ties=efron rl; strata idu;
    weight fw; baseline out=weight survival=s lower=lcl upper=ucl cumhaz=ch
LOGLOGS=logch ;
run;

/*standardized figures*/
data weight; set weight; by idu; output; if last.idu then do; out=10; output; end;
data weight;
set weight;
label idu='ID use';

```

```

format idu yes1fmt.;
proc print data=weight noobs; where out=10; var idu s; title1 "10-year weighted AIDS-free
survival by idu";
goptions reset=all device=zpng ftext="Albany AMT" htext=12pt gsfname=grafout
gsfmode=replace xmax=4 ymax=4 xpixels=4000 ypixels=4000; *1000dpi;
filename grafout "H:\Survival Tutorial\WIHS\standardized.png";
axis1 label=(angle=90 "Probability of AIDS free survival") order=(0 to 1 by .2) w=8 major=(w=8
h=15) minor=none offset=(0,0);
axis2 label=("Years") order=(0 to 10 by 2) w=8 major=(w=8 h=7) minor=none offset=(0,0);
symbol1 c=black v=none i=stepjs w=12 l=1;
symbol2 c=black v=none i=stepjs w=12 l=22;
Legend1 frame
  across=1 down=2
  position = (bottom left inside)
  offset = (10 pct,10 pct) shape=line(1 cm);
proc gplot data=weight; plot s*out=idu/vaxis=axis1 haxis=axis2 nolegend noframe; run; quit;

```

```

goptions reset=all device=zpng ftext="Albany AMT" htext=12pt gsfname=grafout
gsfmode=replace xmax=4 ymax=4 xpixels=4000 ypixels=4000; *1000dpi;
filename grafout "H:\Survival Tutorial\WIHS\standardized_logch.png";
axis1 label=(angle=90 "Log cumulative hazard") order=(-6 to 0 by 0.5) w=8 major=(w=8 h=15)
minor=none offset=(0,0);
axis2 label=("Years") order=(0 to 10 by 2) w=8 major=(w=8 h=7) minor=none offset=(0,0);
symbol1 c=black v=none i=stepjs w=12 l=1;
symbol2 c=black v=none i=stepjs w=12 l=22;
Legend1 frame
  across=1 down=2
  position = (bottom right outside)
  offset = (2 pct,2 pct) shape=line(1 cm);
proc gplot data=weight; plot logch*out=idu/vaxis=axis1 haxis=axis2 nolegend noframe; run;
quit;

```

```

*overall estimate of 10 year risk;
proc phreg data=b; weight fw; model out*yy(0)=/entry=in ties=efron rl; title1 "IPW Cox model";
  baseline out=overallw survival=s lower=lcl upper=ucl;
data overallw2;
set overallw;
by out;
risk=1-s;
riskl=1-ucl;
risku=1-lcl;
if out=9.9274 then do; out=10; output; end;
proc print data=overallw2;
run;

```

```

/*get risk difference and risk ratio at 10 years*/
/*standardized*/

data noidu_r;
set weight;
where out=10 and idu=0;

data idu_r;
set weight;
where out=10 and idu=1;
/*Ten year RD and risk by IDU*/
data tenyr_r;
merge noidu_r (keep=idu s lcl ucl out rename=(s=S_no lcl=lcl_no ucl=ucl_no))
      idu_r (keep= idu s lcl ucl out rename=(s=S_yes lcl=lcl_yes ucl=ucl_yes));
by out;
R_no=1-S_no;
Rnol=1-ucl_no;
Rnou=1-lcl_no;
R_yes=1-S_yes;
Ryesl=1-ucl_yes;
Ryesu=1-lcl_yes;
RD=R_yes-R_no;
drop idu out;
run;
proc print data=tenyr_r;
run;

*Text only - regression-adjusted Cox model;
*without time-varying cov;
proc phreg data=a; model t*y(0)=idu aa age age1-age3 cd4n cd4n1-cd4n3
      /ties=efron rl; title1 "Regression-adjusted Cox model";
      ods select modelinfo parameterestimates fitstatistics;
run;

```