predcumi: A postestimation command for predicting and visualising cumulative incidence estimates after Cox regression models.

Stephen Kaptoge
Department of Public Health and Primary Care
University of Cambridge
Cambridge, UK

18th London Stata Users Group Meeting, Cass Business School, London (13 - 14 September 2012)

Outline

- Background.
  - What motivated this work.
  - Competing risks: Survival vs. cumulative incidence function (CIF).
  - Cause-specific hazard formulation of competing risk models.
  - Competing risk models in Stata (what was missing).

- Implementation of **predcumi**
  - Syntax.
  - Programming aspects.

- Examples: CVD risk prediction
  - Cumulative incidence over observed times vs. within landmark times.
  - Graphical visualisation of the unadjusted/adjusted probabilities.
  - Comparison of covariate adjusted CIF vs. previous implementations.

- Remarks.

Motivation

- Involved in meta-analysis of individual participant data from multiple studies in cardiovascular disease (CVD).
  - CVD is a composite endpoint, comprising CHD and Stroke (each with further subtypes).
  - Aetiological associations suggest some risk factors (e.g. lipids) have stronger associations with CHD and others (e.g. blood pressure) have stronger associations with stroke.
  - Yet it is somewhat traditional to analyse the composite CVD endpoint for risk prediction purposes.

- Question: Could disaggregation of prediction models for CVD to CHD and stroke components possibly lead to better predictions for the composite CVD outcome?
  - What about adjustment for other competing risks (e.g. non-CVD death)?

Competing risks (CR)

- Competing risks (e.g. death from other cause) prevent the event of interest from occurring altogether.
  - Hence, CRs should be handled differently from censoring when interest is on making absolute risk predictions.
  - Otherwise will overestimate the absolute risk of the event of interest.

- Kaplan-Meier failure estimates (**sts generate**) or absolute risk predictions from Cox models (**stcox**) do not account for competing risks …
  - But the **stcox** model provides cause-specific hazards (CSH), the fundamental quantities needed for calculating cumulative incidence to account for competing risks.

CSH, survival, and cumulative incidence

- Cause-specific hazard (CSH) $h_k(t)$: the hazard of failing from a specific cause (k) in the presence of competing risks.

$$h_k(t) = \lim_{\Delta t \to 0} \frac{\text{Prob}(t \le T < t + \Delta t, D = k \mid T \ge t)}{\Delta t}$$

- Cumulative hazard $H_k(t)$ and Survival $S_k(t)$, $S(t)$:

$$H_k(t) = \sum_{j:t_j \le t} h_k(t_j) \qquad S_k(t) = \exp\left(-H_k(t)\right) \qquad S(t) = \exp\left(-\sum_{k=1}^{K} H_k(t)\right)$$

- Cumulative incidence function (CIF) $I_k(t)$:

Overall survival

$$I_k(t) = \sum_{j:t_j \le t} h_k(t_j) S(t_{j-1})$$

- $h_k(t)$ is obtained after **stcox** using --**predict, basehc**-- followed by careful sorting and summations by strata.

Putter H et al *Statist Med* 2007(26)2389–2430

Back to motivating problem …

- Planned to disaggregate the prediction of absolute risk of CVD to CHD and stroke components, treating both as competing events.
  - Calculate CIF of the composite CVD endpoint as sum of the predicted CIFs for CHD and stroke.
  - Additionally treat non-CVD deaths as competing risks.
  - Compare the predictions from the above approaches by calculating discrimination measures (e.g. C-index).

- Needed a program that could calculate the covariate-adjusted CIF over observed failure times, as well as the maximum within user-specified landmark times (e.g. 10-years) for each individual in the dataset.

Competing risk models in Stata (what was missing)

- User-written programs by Enzo Coviello (**stcompet** and **stcompadj**), could only provide CIFs over time adjusted to specific covariate values.

- **stcrreg** was formulated differently than the **stcox** model used for inference of hazard ratios (and was also difficult to converge in large datasets).

- Solution: Write a bespoke postestimation program that predicts CIFs after stcox model (**predcumi**), including graphical visualisation and optional adjustment of covariates.

--predcumi-- Implementation

- Assumes fitted Cox model is for event of interest and gets the specification from **e(***cmdline***)**.

- User provides specification of the competing event(s) via options **compete(***string***)** or **compvars(***varlist***)** and **compcodes(***numlist***)**.

- **--predcumi--** then refits the model for each endpoint to obtain the cause-specific baseline hazards, cause-specific linear predictor, overall baseline survival, and overall linear predictor.

- It then summates $h_k(t_j)$*exp(**xb**$_k$)*$S_0(t_{j-1})$^exp(**xb**) product over time by strata to obtain the CIF for each distinct covariate pattern.
  - This summation part implemented in **mata**, hence faster than adofile.

- If **attimes(***numlist***)** option is specified, it calculates the maximum CIF over each landmark time in **attimes()**.

--predcumi-- Syntax

predcumi [*if exp***] [***in range***] [, compvars(***varlist***) compcodes(***numlist***) compete(***string***) attimes(***numlist integer >0***) lifetimes(***numlist integer >0***) nograph nodots adjust(***string***) xvar(***varname***) compare(***varlist***) nobaseline nobytime** *graph_options* **]**

- Currently uses default output variable names xb*, s0*, hc*, cumhc*, pf*, cif* where * = 1, 2, 3, … endpoints, to save the predictions, but will make flexible in the future.

- **adjust(**string**)** options can be either #, **mean**, **p1**, **p5**, **p10**, **p25**, **p50**, **p75**, **p90**, **p95**, **p99**, or **round(#),** in which case the adjusted CIF calculations are done after fixing the covariates at the values specified.

Example: CVD risk prediction

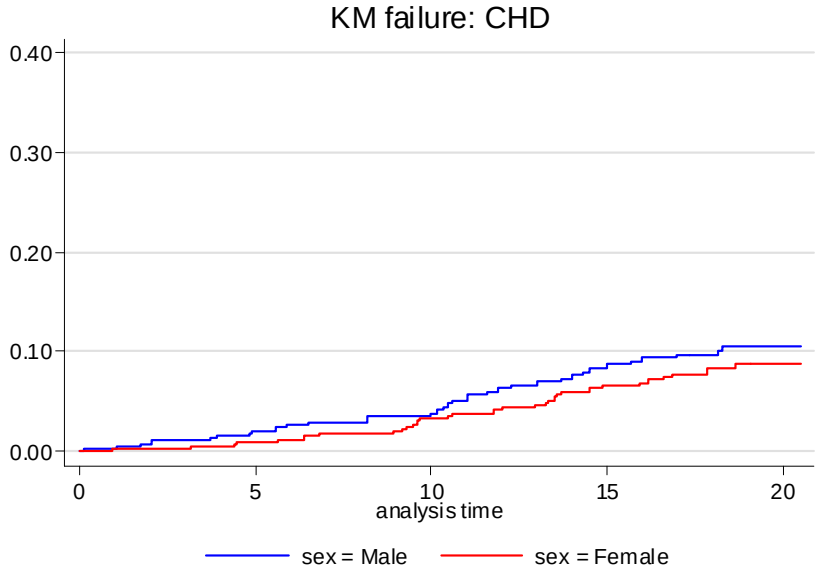
```
Contains data from data\predcumi_demo_data.dta
  obs:           817
 vars:            15                          11 Sep 2012 15:28
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
cohort          str7   %9s                    Cohort abbreviation
subjectid       str3   %9s                    Subject ID
sex             byte   %8.0g       sex        Sex
duration        float  %9.0g                  Time to event/censoring (yrs)
ep_chdmi        byte   %9.0g                  CHD
ep_crbv         byte   %8.0g                  Stroke
ep_cvd          byte   %9.0g                  CVD (CHD or stroke)
ep_ncv_f        byte   %9.0g                  Non-CVD death
ep_dead         byte   %9.0g                  Any death
ages            float  %9.0g                  Age at survey (yrs)
smallbin        byte   %9.0g       statbin    Smoking status
sbp             int    %8.0g                  SBP (mmHg)
hxdiabbin       byte   %17.0g      hx         History of diabetes
tchol           float  %9.0g                  Total cholesterol (mmol/l)
hdl             float  %9.0g                  HDL-C (mmol/l)
-------------------------------------------------------------------------------
Sorted by:  cohort  subjectid
```

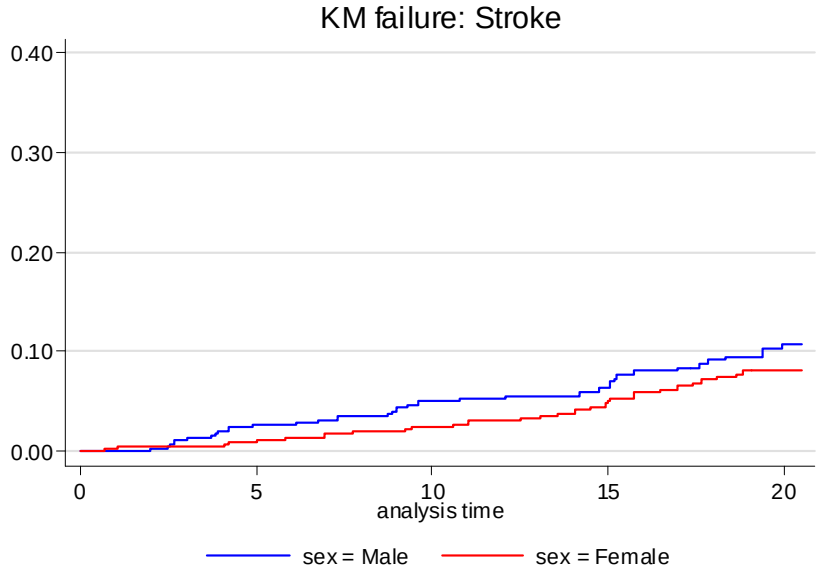
Data descriptive statistics

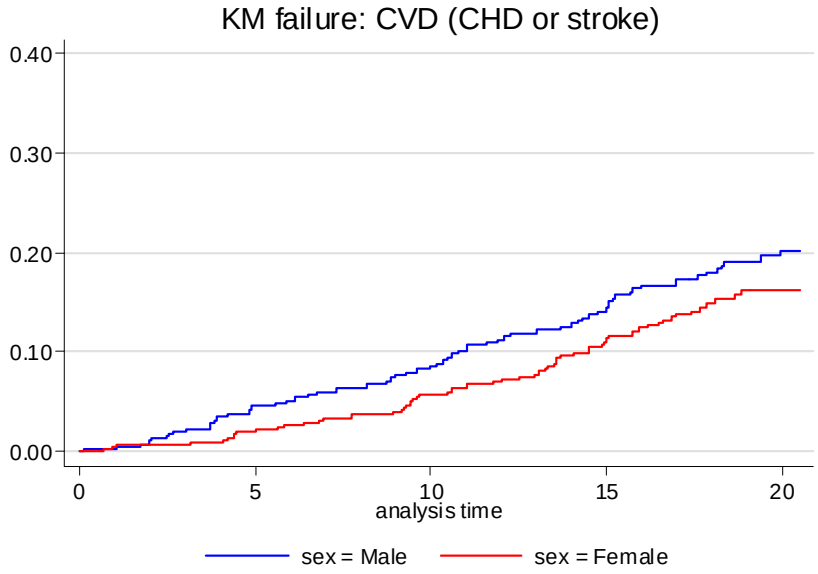
Variable	Obs	Mean	Std. Dev.	Min	Max
cohort	0				
subjectid	0				
sex	817	1.512852	.500141	1	2
sexbin	817	.5128519	.500141	0	1
duration	817	16.74673	5.693721	.0821355	20.4627
ages	817	57.85982	11.38797	40	80.096
smallbin	817	.2423501	.4287675	0	1
sbp	817	145.1004	21.50535	91	230
hxdiabbin	817	.0330477	.1788707	0	1
tchol	817	5.727442	1.028323	3.08	10.24
hdl	817	1.473035	.3640861	.7	2.97

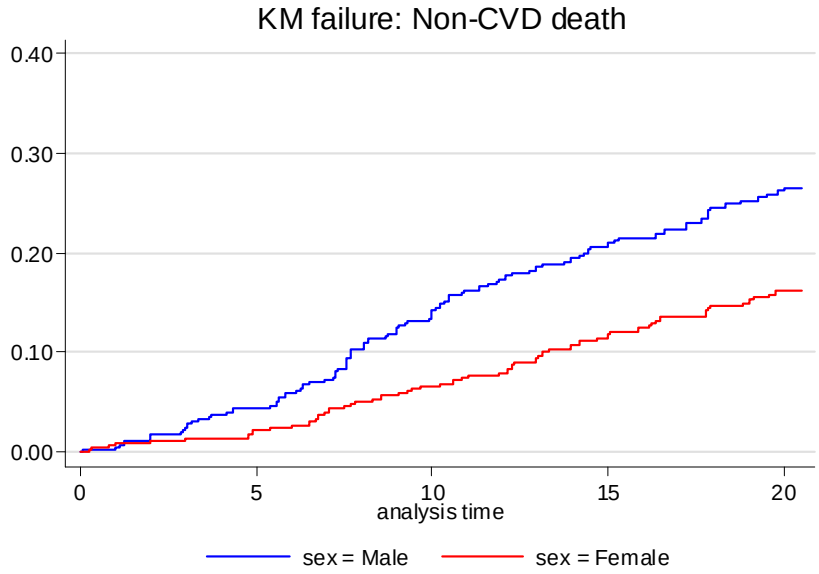
variable	N	sum	
ep_chdmi	817	66	CHD
ep_crbv	817	63	Stroke
ep_cvd	817	129	CVD (CHD or stroke)
ep_ncv_f	817	157	Non-CVD death
ep_dead	817	240	Any death

KM failure estimates









User-fitted CHD model (event of interest)

```
. stset duration, failure("ep_chdmi == 1")
. xi: stcox ages smallbin sbp hxdiabbin tchol hdl, strata(sex) nolog
         failure _d:  ep_chdmi == 1
   analysis time _t:  duration
Stratified Cox regr. -- Breslow method for ties
No. of subjects =          817                     Number of obs   =       817
No. of failures =           66
Time at risk    =  13682.08051
                                                    LR chi2(6)      =     89.58
Log likelihood  =   -338.07286                      Prob > chi2     =    0.0000
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ages |   1.085469   .0150106     5.93   0.000     1.056444    1.115292
    smallbin |   2.101201   .6167235     2.53   0.011     1.182038    3.735113
         sbp |   1.023793   .0060849     3.96   0.000     1.011936    1.035788
   hxdiabbin |   1.883857   .8553814     1.39   0.163      .773664    4.587153
       tchol |   1.041145    .128146     0.33   0.743     .8179821    1.325191
         hdl |    1.17283   .3919691     0.48   0.633     .6091954    2.257945
------------------------------------------------------------------------------
                                                              Stratified by sex
```

--predcumi-- Model 1 (CHD)

```
. predcumi, compvars("ep_crbv ep_ncv_f") compcodes(1) attimes(10)
        failure _d:  ep_chdmi == 1
  analysis time _t:  duration
Stratified Cox regr. -- Breslow method for ties
No. of subjects =          817                     Number of obs   =       817
No. of failures =           66
Time at risk    =  13682.08051
                                                    LR chi2(6)      =     89.58
Log likelihood  =   -338.07286                      Prob > chi2     =    0.0000
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ages |   1.085469   .0150106     5.93   0.000     1.056444    1.115292
    smallbin |   2.101201   .6167235     2.53   0.011     1.182038    3.735113
         sbp |   1.023793   .0060849     3.96   0.000     1.011936    1.035788
   hxdiabbin |   1.883857   .8553814     1.39   0.163      .773664    4.587153
       tchol |   1.041145    .128146     0.33   0.743     .8179821    1.325191
         hdl |    1.17283   .3919691     0.48   0.633     .6091954    2.257945
------------------------------------------------------------------------------
                                                        Stratified by sex

Duplicates in terms of sex _t0 _t hc1
      dupl1 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         60       90.91       90.91
          1 |          6        9.09      100.00
------------+-----------------------------------
      Total |         66      100.00
```

--predcumi-- Model 2 (Stroke)

```
-> stset duration, failure(ep_crbv==1)
         failure _d:  ep_crbv == 1
   analysis time _t:  duration
Stratified Cox regr. -- Breslow method for ties
No. of subjects =          817                     Number of obs   =       817
No. of failures =           63
Time at risk    =  13682.08051
                                                    LR chi2(6)      =     45.67
Log likelihood  =   -341.26046                      Prob > chi2     =    0.0000
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ages |   1.080394   .0146957     5.68   0.000     1.051971    1.109584
    smallbin |   1.239853   .3939733     0.68   0.499     .6651106    2.311247
         sbp |   1.004778   .0065243     0.73   0.463     .9920712    1.017647
   hxdiabbin |   1.123873   .6831708     0.19   0.848      .341425    3.699464
       tchol |   .9629966   .1232594    -0.29   0.768     .7493327    1.237584
         hdl |   .6935387   .2667539    -0.95   0.341     .3263438    1.473893
------------------------------------------------------------------------------
                                                              Stratified by sex

Duplicates in terms of sex _t0 _t hc2
      dupl2 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         57       90.48       90.48
          1 |          6        9.52      100.00
------------+-----------------------------------
      Total |         63      100.00
```

--predcumi-- Model 3 (Non-CVD death)

```
-> stset duration, failure(ep_ncv_f==1)
         failure _d:  ep_ncv_f == 1
   analysis time _t:  duration
Stratified Cox regr. -- Breslow method for ties
No. of subjects =          817                     Number of obs   =       817
No. of failures =          157
Time at risk    =  13682.08051
                                                   LR chi2(6)      =    175.09
Log likelihood  =   -822.29807                     Prob > chi2     =    0.0000
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ages |   1.105477   .0099111    11.18   0.000     1.086221    1.125074
    smallbin |   1.661562   .3161485     2.67   0.008     1.144344    2.412549
         sbp |   .9982142   .0041068    -0.43   0.664     .9901974    1.006296
   hxdiabbin |   1.981136   .6249526     2.17   0.030     1.067583    3.676438
       tchol |   .8962753   .0760461    -1.29   0.197     .7589614    1.058432
         hdl |   1.454097   .3178565     1.71   0.087     .9473837    2.231828
------------------------------------------------------------------------------
                                                              Stratified by sex

Duplicates in terms of sex _t0 _t hc3
      dupl3 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        109       69.43       69.43
          1 |         32       20.38       89.81
          2 |         12        7.64       97.45
          3 |          4        2.55      100.00
------------+-----------------------------------
      Total |        157      100.00
```

--predcumi-- Model 4 (Any event)

```
-> stset duration, failure(__000002)
        failure _d:  __000002
  analysis time _t:  duration
Stratified Cox regr. -- Breslow method for ties
No. of subjects =          817                     Number of obs   =       817
No. of failures =          286
Time at risk    =  13682.08051
                                                   LR chi2(6)      =    290.60
Log likelihood  =   -1511.4978                     Prob > chi2     =    0.0000
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ages |   1.094803   .0071982    13.78   0.000     1.080785    1.109002
    smallbin |   1.636837   .2329696     3.46   0.001     1.238382    2.163496
         sbp |   1.006025   .0030151     2.00   0.045     1.000133    1.011952
   hxdiabbin |   1.789991   .4242442     2.46   0.014     1.124885     2.84835
       tchol |   .9462233   .0580374    -0.90   0.367     .8390435    1.067094
         hdl |   1.187661   .1965384     1.04   0.299     .8586825    1.642678
------------------------------------------------------------------------------
                                                              Stratified by sex

Duplicates in terms of sex _t0 _t hc
       dupl |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        144       50.35       50.35
          1 |         94       32.87       83.22
          2 |         36       12.59       95.80
          3 |         12        4.20      100.00
------------+-----------------------------------
      Total |        286      100.00
```

--predcumi-- Model 1 CIF over time

```
Calculating cumulative incidence of ep_chdmi over time among 817 subjects
in 2 strata at 817 linear predictor values
Stratum = 1 (Male), n = 398, and  398 linear predictors
..................................................    50
..................................................   100
..................................................   150
..................................................   200
..................................................   250
..................................................   300
..................................................   350
..............................................

Stratum = 2 (Female), n = 419, and  419 linear predictors
..................................................    50
..................................................   100
..................................................   150
..................................................   200
..................................................   250
..................................................   300
..................................................   350
..................................................   400
...................
```

--predcumi-- Model 1 CIF up to time 10

```
Calculating cumulative incidence of ep_chdmi at time = 10
in 2 strata
Stratum = 1 (Male), n = 398, and 398 linear predictors
..................................................    50
..................................................   100
..................................................   150
..................................................   200
..................................................   250
..................................................   300
..................................................   350
..........................................

Stratum = 2 (Female), n = 419, and 419 linear predictors
..................................................    50
..................................................   100
..................................................   150
..................................................   200
..................................................   250
..................................................   300
..................................................   350
..................................................   400
..................
```

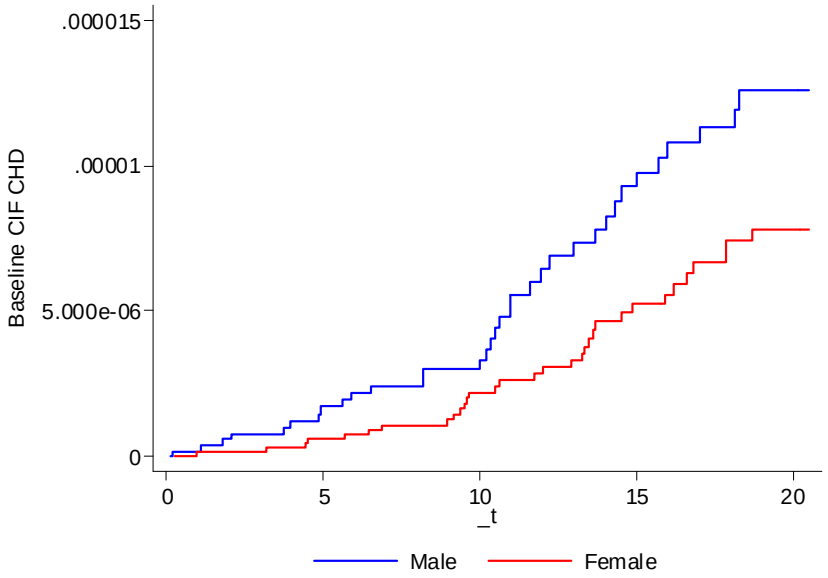
… and similarly for each of Model 2 (Stroke) and Model 3 (Non-CVD death)

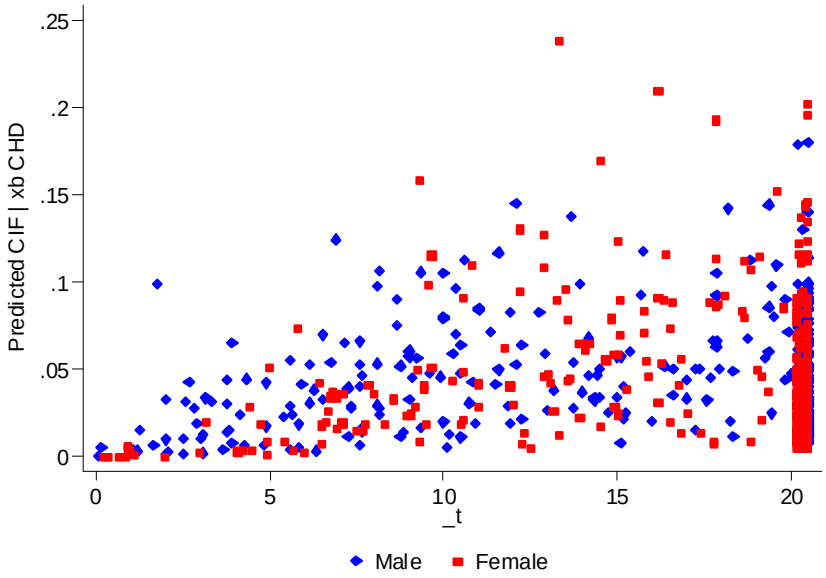
Execution time = 4 sec for calculations + 12 sec for graphs

--predcumi-- Output variables (default)

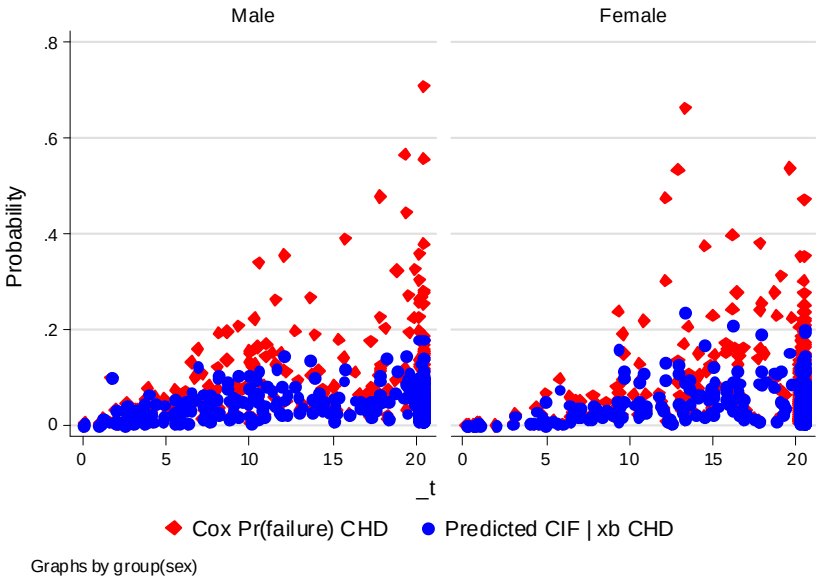
```
Generated variables
              storage  display     value
variable name   type   format      label      variable label
------------------------------------------------------------------------------
xb1             double %10.0g                 Linear prediction CHD
xb2             double %10.0g                 Linear prediction Stroke
xb3             double %10.0g                 Linear prediction Non-CVD death
xb              double %10.0g                 Linear prediction overall
cif01           double %10.0g                 Baseline CIF CHD
cif1            double %10.0g                 Predicted CIF | xb CHD
cif01t10        double %10.0g                 Baseline CIF (T <= 10) CHD
cif1t10         double %10.0g                 Predicted CIF (T <= 10) CHD
cif02           double %10.0g                 Baseline CIF Stroke
cif2            double %10.0g                 Predicted CIF | xb Stroke
cif02t10        double %10.0g                 Baseline CIF (T <= 10) Stroke
cif2t10         double %10.0g                 Predicted CIF (T <= 10) Stroke
cif03           double %10.0g                 Baseline CIF Non-CVD death
cif3            double %10.0g                 Predicted CIF | xb Non-CVD death
cif03t10        double %10.0g                 Baseline CIF (T <= 10) Non-CVD death
cif3t10         double %10.0g                 Predicted CIF (T <= 10) Non-CVD death
pf1             double %10.0g                 Cox Pr(failure) CHD
pf1t10          double %10.0g                 Cox Pr(T <= 10) CHD
pf2             double %10.0g                 Cox Pr(failure) Stroke
pf2t10          double %10.0g                 Cox Pr(T <= 10) Stroke
pf3             double %10.0g                 Cox Pr(failure) Non-CVD death
pf3t10          double %10.0g                 Cox Pr(T <= 10) Non-CVD death
pf              double %10.0g                 Cox Pr(failure) overall
pft10           double %10.0g                 Cox Pr(T <= 10) overall
```

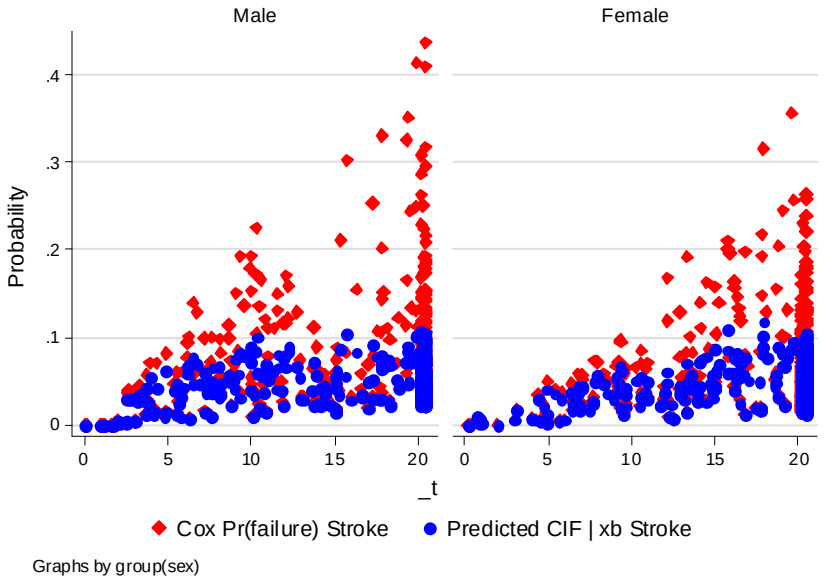
Baseline CIF and predicted CIF vs. time

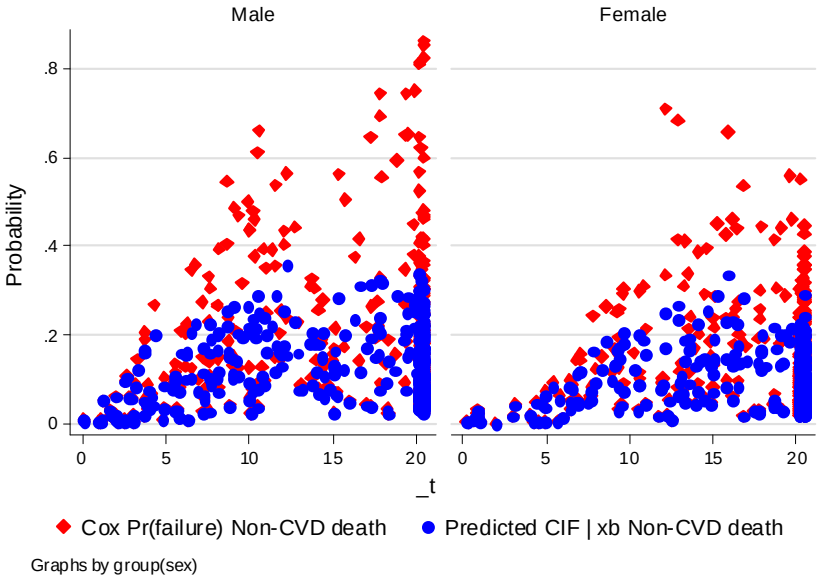




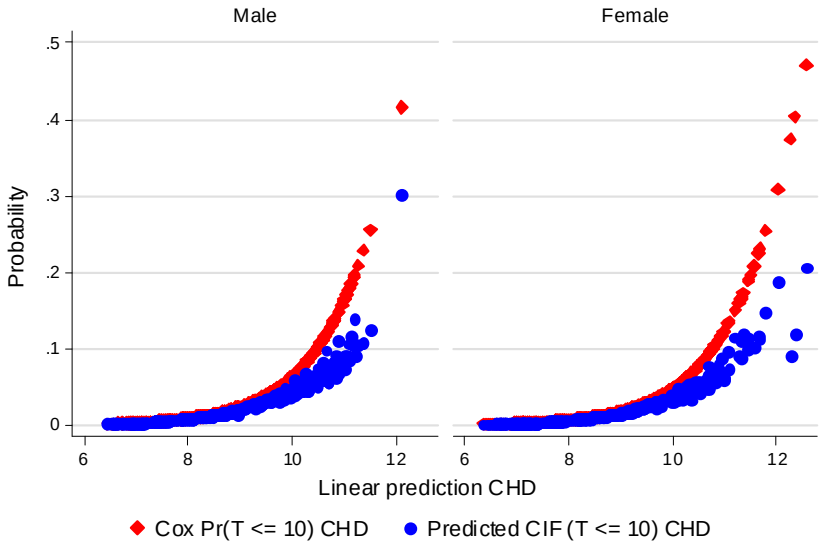
Cox Pr(failure) and predicted CIF vs. time

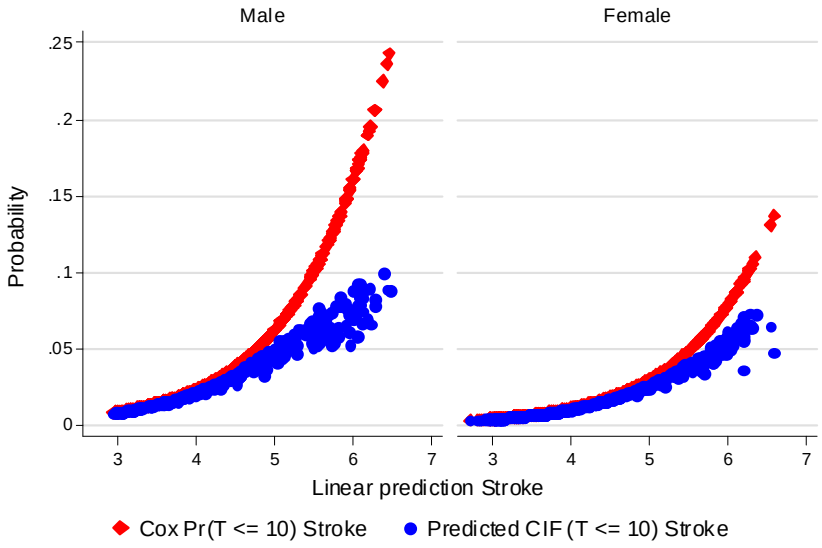


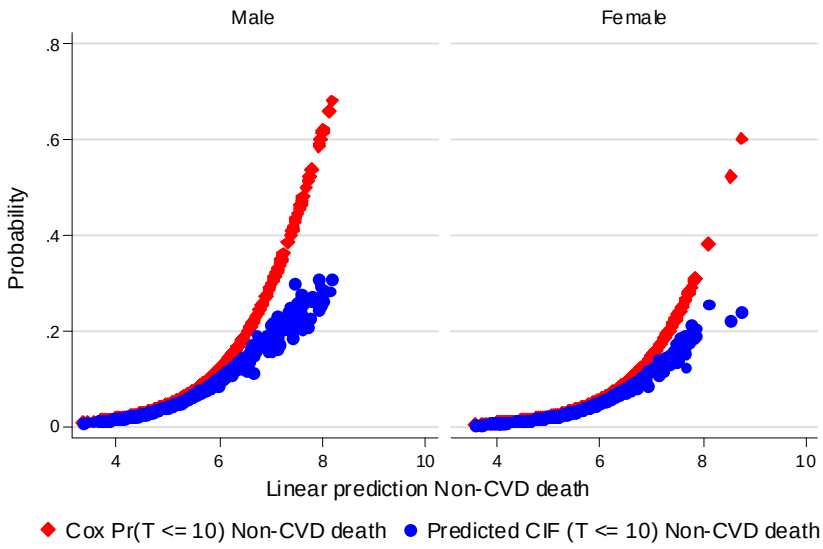




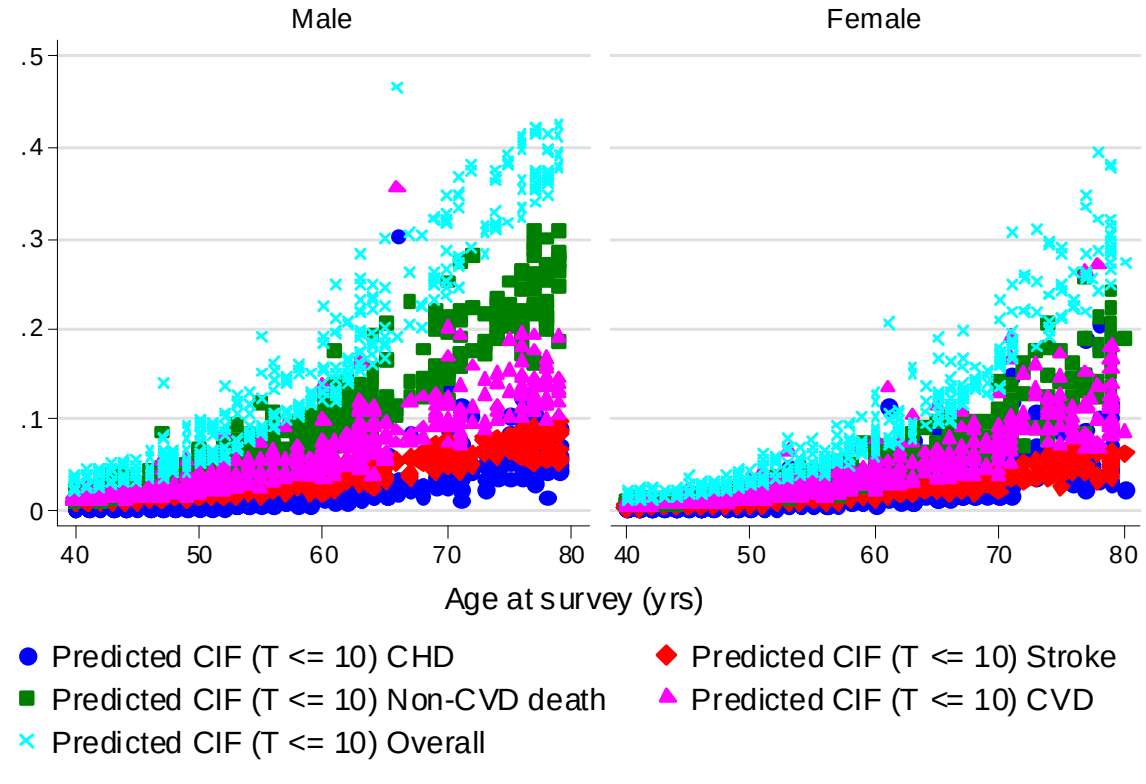
Cox Pr(failure) and predicted CIF vs. linear predictor







DIY combined summary of the predicted CIFs vs. age



Where: $CIF_{CVD} = CIF_{CHD} + CIF_{Stroke}$

$CIF_{Overall} = CIF_{CHD} + CIF_{Stroke} + CIF_{Non\text{-}CVD}$

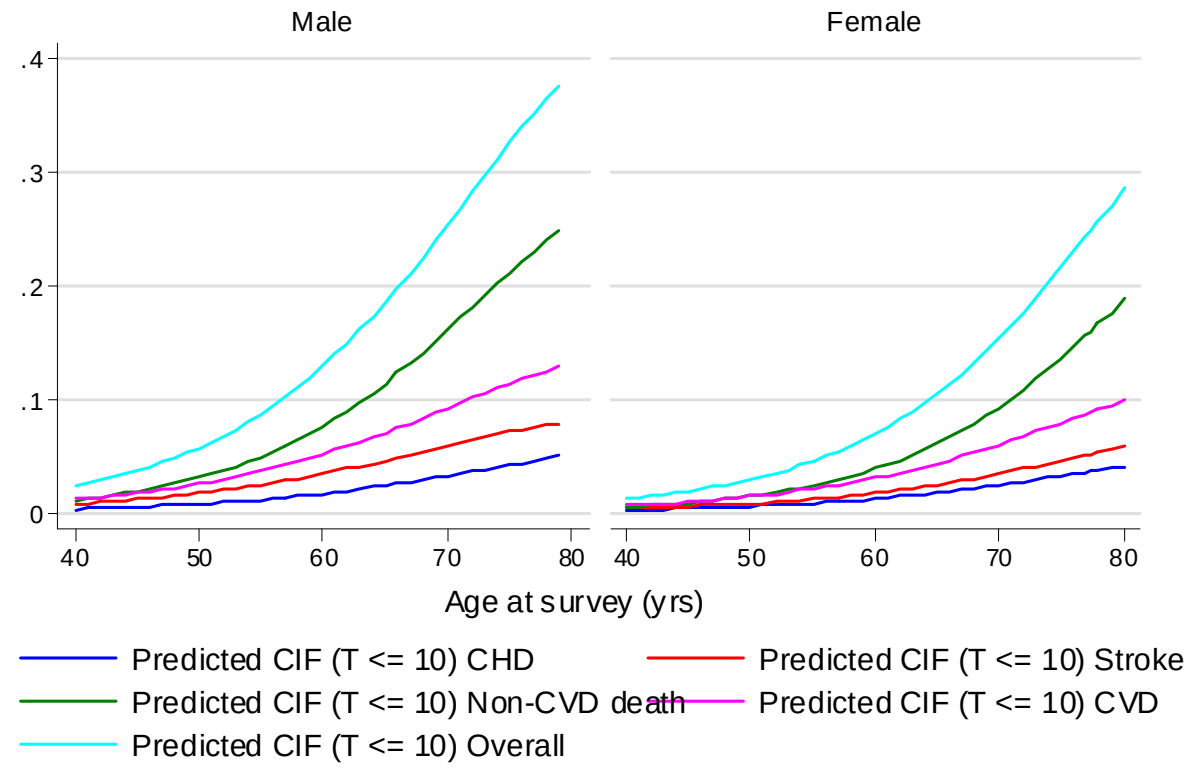
--predcumi-- Covariate adjusted CIF (1)

```
. predcumi, compvars("ep_crbv ep_ncv_f") compcodes(1) attimes(10) ///
  adjust(smallbin = 0 sbp = mean hxdiabbin = 0 tchol = mean hdl = mean) ///
  xvar(ages) nobaseline nobytime

< output omitted, same models as described before >

Adjusting covariates to specified values ...
Adjusting smallbin            = 0
Adjusting sbp                 = 145.1003671970624     (mean)
Adjusting hxdiabbin           = 0
Adjusting tchol               = 5.727441867678956     (mean)
Adjusting hdl                 = 1.473035494437854     (mean)

Unadjusted variables ages
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        ages |       817    57.85982    11.38797         40     80.096

Calculating cumulative incidence of ep_chdmi at time = 10
in 2 strata
Stratum = 1 (Male), n = 398, and 40 linear predictors
.........................................
Stratum = 2 (Female), n = 419, and 42 linear predictors
.........................................
```

… and similarly for each of Model 2 (Stroke) and Model 3 (Non-CVD death)

Execution time = 1 sec for calculations + 6 sec for graphs

DIY combined summary of the predicted CIFs vs. age



Where: $CIF_{CVD} = CIF_{CHD} + CIF_{Stroke}$

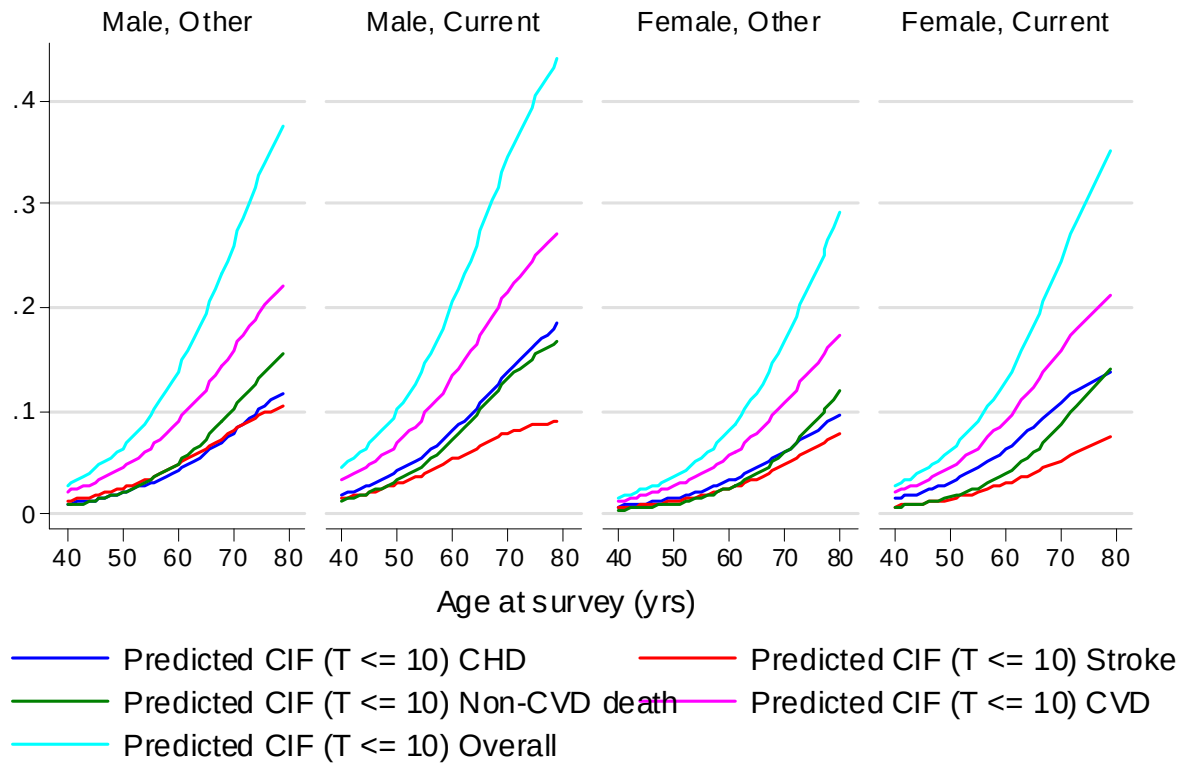
$CIF_{Overall} = CIF_{CHD} + CIF_{Stroke} + CIF_{Non\text{-}CVD}$

--predcumi-- Covariate adjusted CIF (2)

```
. predcumi, compvars("ep_crbv ep_ncv_f") compcodes(1) attimes(10) ///
  adjust(sbp = p95 hxdiabbin = 0 tchol = p95 hdl = p5) ///
  xvar(ages) nobaseline nobytime compare(smallbin)

< output omitted, same models as described before >

Adjusting covariates to specified values ...
Adjusting sbp                 = 183                  (p95)
Adjusting hxdiabbin           = 0
Adjusting tchol               = 7.46999979019165     (p95)
Adjusting hdl                 = .9599999785423279    (p5)
Unadjusted variables ages smallbin

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        ages |       817    57.85982    11.38797         40     80.096
    smallbin |       817    .2423501    .4287675          0          1

Calculating cumulative incidence of ep_chdmi at time = 10
in 2 strata
Stratum = 1 (Male), n = 398, and 78 linear predictors
..................................................    50
..........................
Stratum = 2 (Female), n = 419, and 71 linear predictors
..................................................    50
..................
```

... and similarly for each of Model 2 (Stroke) and Model 3 (Non-CVD death)

DIY combined summary of the predicted CIFs vs. age

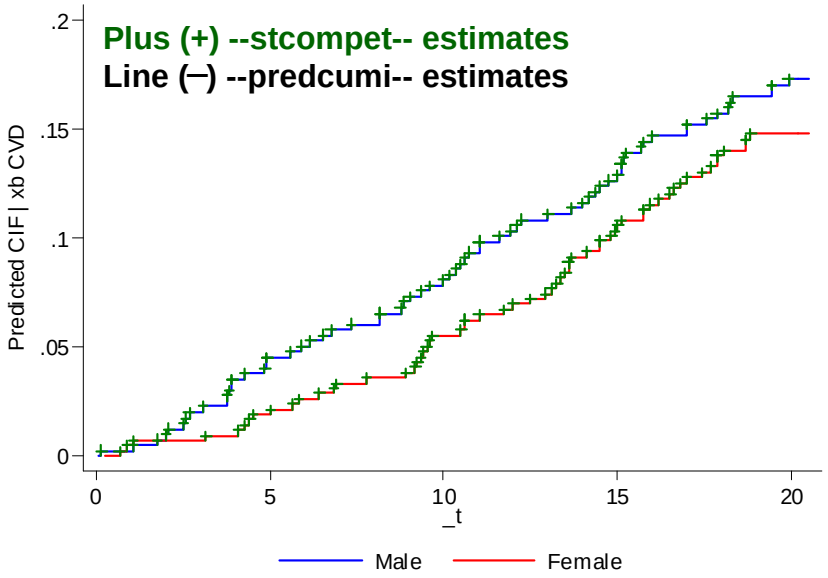


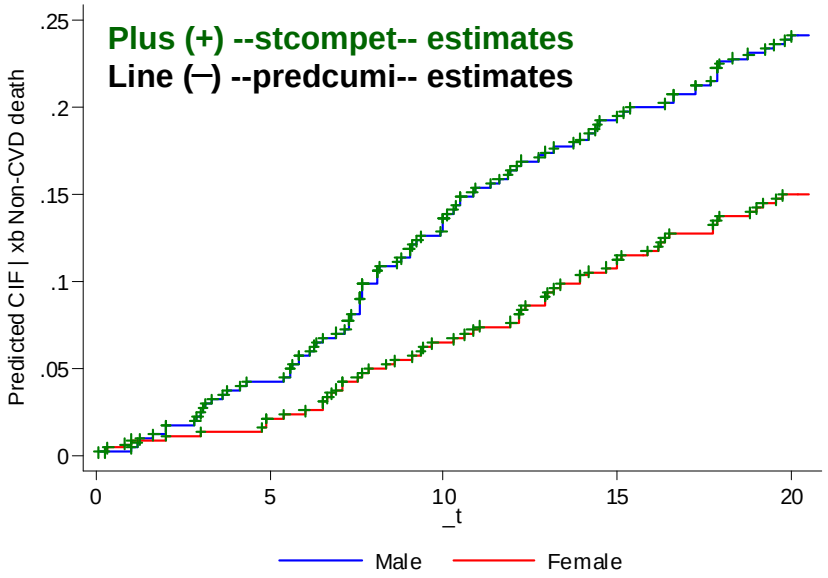
Where: $CIF_{CVD} = CIF_{CHD} + CIF_{Stroke}$

$CIF_{Overall} = CIF_{CHD} + CIF_{Stroke} + CIF_{Non\text{-}CVD}$

Check against --stcompet-- (unadjusted)

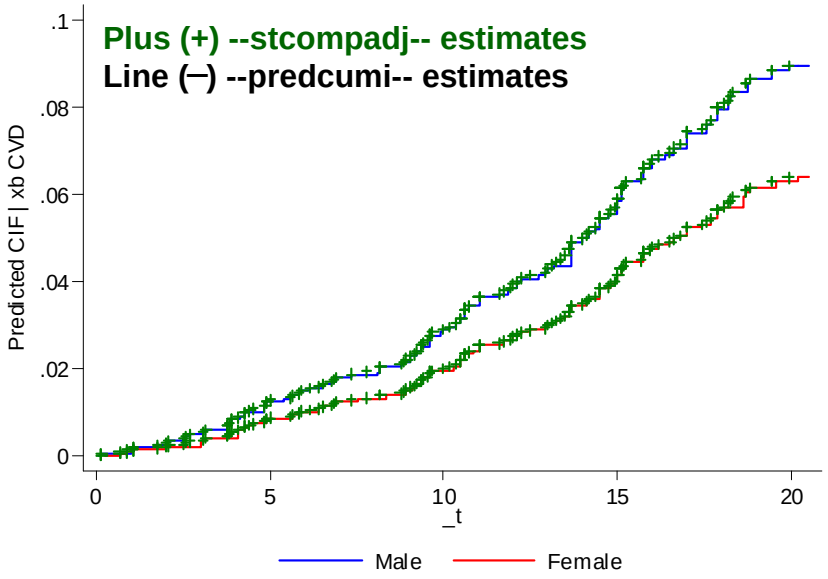
```
. stset duration, failure("cause == 1")
. xi: stcox, strata(sex) estimate nolog
. predcumi, compete("cause == 2")
. stcompet cumi = ci, compet1(2) by(sex)
  foreach m of numlist 1 2 {
    overlay cif`m' _t, over(sex) ms(i i) c(J) ylabel(#5, angle(horiz)) ///
    scheme("s1color_sk") addplot(scatter cumi _t if cause == `m', ms(+))
    name(gr1_cause`m', replace)
  }
```

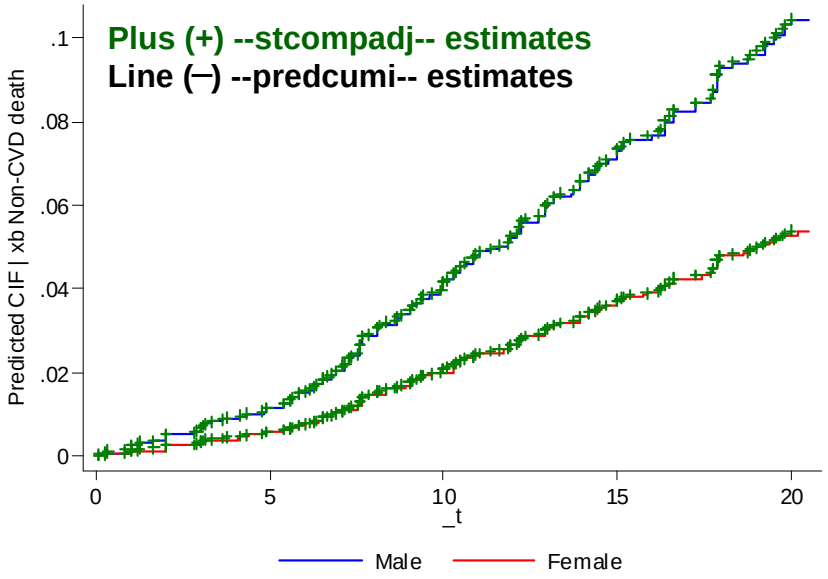




Check against --stcompadj-- (adjusted)

```
. stset duration, failure("cause == 1")
. xi: stcox sexbin ages, nolog
. predcumi, compete("cause==2") adjust(ages = 50)
. local fitopts = "maineffect(sexbin ages) competeffect(sexbin ages) showmod"
. stcompadj sexbin=0 ages = 50, compet(2) gen(cifmain0 cifcomp0) `fitopts'
. stcompadj sexbin=1 ages = 50, compet(2) gen(cifmain1 cifcomp1) `fitopts'
  foreach m of numlist 1 2 {
    local compvars = cond(`m' == 1, "cifmain*", "cifcomp*")
    overlay cif`m' _t, over(sexbin) ms(i i) c(J) ylabel(#5, angle(horiz)) ///
     scheme("s1color_sk") addplot(scatter `compvars' _t if cause == `m', ms(+ +))
    name (gr3_cause`m', replace)
  }
```





Check against --stcrreg-- (adjusted)

```
. stset duration, failure("cause == 1")
. xi: stcox sexbin ages, nolog
        failure _d:  cause == 1
  analysis time _t:  duration
Cox regression -- Breslow method for ties
No. of subjects =          817                     Number of obs   =       817
No. of failures =          129
Time at risk    =  13682.08051
                                                   LR chi2(2)      =    111.93
Log likelihood  =   -780.89034                     Prob > chi2     =    0.0000
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      sexbin |    .681508   .1204439    -2.17   0.030     .4819888    .9636182
        ages |   1.092081   .0096398     9.98   0.000      1.07335    1.111139
------------------------------------------------------------------------------
. predcumi, compete("cause==2") adjust(ages = 50)
<output, omitted >
. foreach m of numlist 1 2 {
    levelsof cause if !inlist(cause, 0, `m'), local(compevent)
    stset duration, failure("cause == `m'")
    xi: stcrreg sexbin ages, compete(cause == `compevent')
    local atval1 = "at1(sexbin = 0 ages = 50) at2(sexbin = 1 ages = 50)"
    forvalues k = 1/1 {
          local gropts = `"title("") ylabel(#5, angle(horiz))"'
          local atvals "`atval`k''"
          local addplot = "scatter cif`m' _t, ms(+)"
          stcurve, cif `atvals' `gropts' addplot(`addplot') ///
          legend(order(1 2)) name(gr4at`k'_cause`m', replace)
    }
 }
```

Check against --stcrreg-- Model 1

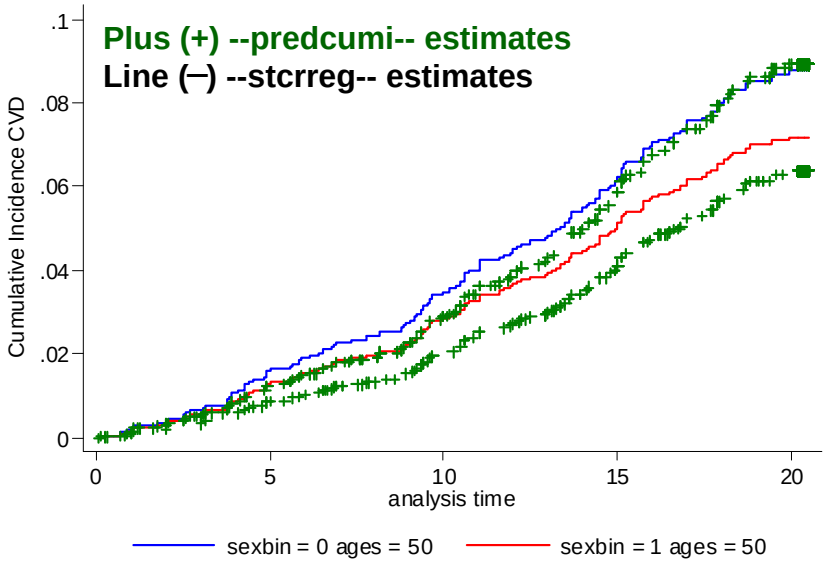
```
 failure _d:  cause == 1
   analysis time _t:  duration

Iteration 0:   log pseudolikelihood = -821.17236
Iteration 1:   log pseudolikelihood = -817.50996
Iteration 2:   log pseudolikelihood = -817.48958
Iteration 3:   log pseudolikelihood = -817.48958

Competing-risks regression                          No. of obs       =       817
                                                    No. of subjects  =       817
Failure event  : cause == 1                         No. failed       =       129
Competing event: cause == 2                         No. competing    =       157
                                                    No. censored     =       531
                                                    Wald chi2(2)     =     65.94
Log pseudolikelihood = -817.48958                   Prob > chi2      =    0.0000

------------------------------------------------------------------------------
             |               Robust
          _t |        SHR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      sexbin |   .8124038    .143884    -1.17   0.241     .5741383    1.149549
        ages |   1.068916   .0090733     7.85   0.000      1.05128    1.086848
------------------------------------------------------------------------------
```

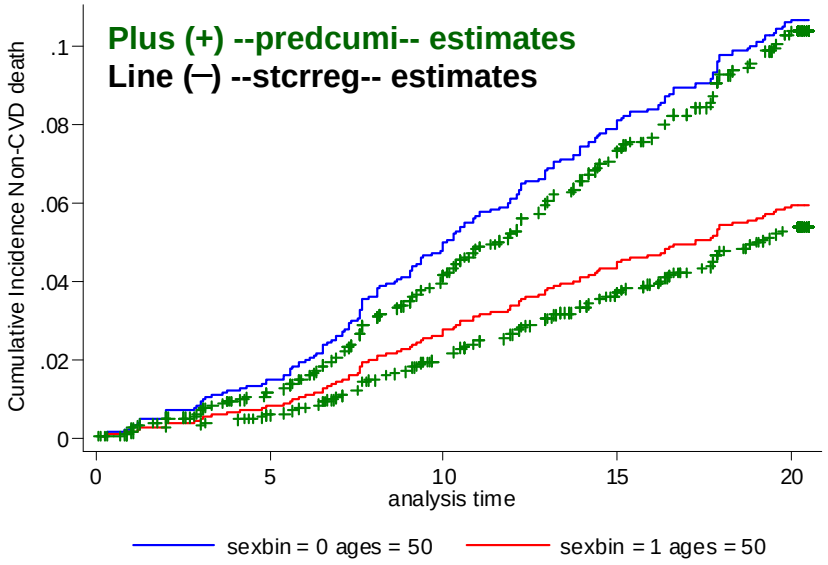
Check against --stcrreg-- Model 2

```
 failure _d:  cause == 2
   analysis time _t:  duration

Iteration 0:   log pseudolikelihood = -971.95067
Iteration 1:   log pseudolikelihood = -969.36457
Iteration 2:   log pseudolikelihood =  -969.3526
Iteration 3:   log pseudolikelihood =  -969.3526

Competing-risks regression                          No. of obs       =       817
                                                    No. of subjects  =       817
Failure event  : cause == 2                         No. failed       =       157
Competing event: cause == 1                         No. competing    =       129
                                                    No. censored     =       531
                                                    Wald chi2(2)     =    117.48
Log pseudolikelihood =  -969.3526                   Prob > chi2      =    0.0000

------------------------------------------------------------------------------
             |               Robust
          _t |        SHR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      sexbin |   .5422991   .0886548    -3.74   0.000      .393627    .7471242
        ages |    1.08355   .0084414    10.30   0.000     1.067131    1.100222
------------------------------------------------------------------------------
```

Check against --stcrreg--

- Not so good agreement …, but not a surprise considering the model formulation in --**stcrreg**-- is different from --**stcox**--.





Remarks

- --**predcumi**-- provides some added functionality not addressed in previous user-written programs, i.e. individual predictions and CIF within landmark times.
  - Ongoing considerations include extension to lifetime-risk calculation with age-as-timescale models.

- --**stcompet**-- and --**stcompadj**--, however, do a lot more with respect to covariate adjusted CIF estimates, including confidence intervals, hypothesis testing, and CIF inferences based on flexible parametric models.
  - Also discovered during this meeting that **--stpm2--** has a wrapper function **--stpm2cif--** for adjusted CIF calculations.

- For large datasets, competing risk models based on cause-specific hazards could be more tractable than currently possible with --**stcrreg**-- (??)

Acknowledgements

- Mentors
  - S Thompson, J Danesh
- Statistical collaborators
  - S Thompson, I White, L Pennells, A Wood
- ERFC collaborators
  - http://www.phpc.cam.ac.uk/ceu/research/erfc/studies/