

# Issues in Standardization

Paul C Lambert<sup>1,2</sup>

<sup>1</sup>Biostatistics Research Group, Department of Health Sciences, University of Leicester, UK

<sup>2</sup>Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Symposium for statisticians working in register-based cancer  
epidemiology

Stockholm, 26 September 2019



# Age Standardisation

- This talk is about taking averages from predictions from a statistical model, i.e. regression standardization.
- I will concentrate on relative survival models. Methods/software also appropriate for standard and competing risks models.
- I will basically be describing the `standsurv` command.
- Usual practice: Take a weighted average over age groups to get age standardised relative survival.

$$\bar{R}(t) = \sum_{j=1}^5 w_j \hat{R}_j(t)$$

- Usually estimate  $R_j(t)$  non-parametrically (Pohar-Perme / Ederer II).

# Age Standardisation

- We can also age standardize from a model.

$$R_i(t) = f(t|Z_i)$$

$$\bar{R}(t) = \frac{1}{N} \sum_{i=1}^N w_i \hat{R}_i(t|Z_i)$$

- $w_i$  up or down-weights relative to external population.
- I will generalise to covariates  $Z$  with exposure  $X$
- I am essentially using the G-formula, without time-dependent exposures/confounding.

# Example model

- Men diagnosed with Melanoma in England.
- Five deprivation groups. Here I restrict to least and most deprived groups (X).
- Relative survival model
- Model the effect of age continuously using splines (Z).
- Interaction between age and deprivation groups.
- Both age and deprivation group have non-proportional effects.

```
stpm2 dep5 agercs1 agercs2 agercs3 ///  
      agercs1_dep5 agercs2_dep5 agercs3_dep5, ///  
      scale(hazard) df(5) bhazard(rate) ///  
      tvc(dep5 agercs1 agercs2 agercs3) dftvc(3)
```

# Example model

- Men diagnosed with Melanoma in England.
- Five deprivation groups. Here I restrict to least and most deprived groups (X).
- Relative survival model
- Model the effect of age continuously using splines (Z).
- Interaction between age and deprivation groups.
- Both age and deprivation group have non-proportional effects.

```
stpm2 dep5 agercs1 agercs2 agercs3 ///  
      agercs1_dep5 agercs2_dep5 agercs3_dep5, ///  
      scale(hazard) df(5) bhazard(rate) ///  
      tvc(dep5 agercs1 agercs2 agercs3) dftvc(3)
```

# Example model

- Men diagnosed with Melanoma in England.
- Five deprivation groups. Here I restrict to least and most deprived groups (X).
- Relative survival model
- Model the effect of age continuously using splines (Z).
- Interaction between age and deprivation groups.
- Both age and deprivation group have non-proportional effects.

```
stpm2 dep5 agercs1 agercs2 agercs3 ///  
      agercs1_dep5 agercs2_dep5 agercs3_dep5, ///  
      scale(hazard) df(5) bhazard(rate) ///  
      tvc(dep5 agercs1 agercs2 agercs3) dftvc(3)
```

# Example model

- Men diagnosed with Melanoma in England.
- Five deprivation groups. Here I restrict to least and most deprived groups (X).
- Relative survival model
- Model the effect of age continuously using splines (Z).
- Interaction between age and deprivation groups.
- Both age and deprivation group have non-proportional effects.

```
stpm2 dep5 agercs1 agercs2 agercs3 ///  
      agercs1_dep5 agercs2_dep5 agercs3_dep5, ///  
      scale(hazard) df(5) bhazard(rate) ///  
      tvc(dep5 agercs1 agercs2 agercs3) dftvc(3)
```

# Example model

- Men diagnosed with Melanoma in England.
- Five deprivation groups. Here I restrict to least and most deprived groups (X).
- Relative survival model
- Model the effect of age continuously using splines (Z).
- Interaction between age and deprivation groups.
- Both age and deprivation group have non-proportional effects.

```
stpm2 dep5 agercs1 agercs2 agercs3 ///  
      agercs1_dep5 agercs2_dep5 agercs3_dep5, ///  
      scale(hazard) df(5) bhazard(rate) ///  
      tvc(dep5 agercs1 agercs2 agercs3) dftvc(3)
```



# Example model

- Men diagnosed with Melanoma in England.
- Five deprivation groups. Here I restrict to least and most deprived groups (X).
- Relative survival model
- Model the effect of age continuously using splines (Z).
- Interaction between age and deprivation groups.
- Both age and deprivation group have non-proportional effects.

```
stpm2 dep5 agerics1 agerics2 agerics3 ///  
      agerics1_dep5 agerics2_dep5 agerics3_dep5, ///  
      scale(hazard) df(5) bhazard(rate) ///  
      tvc(dep5 agerics1 agerics2 agerics3) dftvc(3)
```

# Internal Age-Standardisation

- Take average within each deprivation group
  - 25,618 in least deprived and 9672 in most deprived.

$$\bar{R}(t|X=x, Z_{X=x}) = \frac{1}{N_{X=x}} \sum_{i=1}^{N_{X=x}} \hat{R}_i(t|X=x, Z_i)$$

# Internal Age-Standardisation

- Take average within each deprivation group
  - 25,618 in least deprived and 9672 in most deprived.

$$\bar{R}(t|X=x, Z_{X=x}) = \frac{1}{N_{X=x}} \sum_{i=1}^{N_{X=x}} \hat{R}_i(t|X=x, Z_i)$$

```
standsurv, survival timevar(tt) ci    ///  
  at1(., atif(dep5==1))              ///  
  at2(., atif(dep5==0))              ///  
  atvar(RS_dep5 RS_dep1)             ///  
  contrast(difference)                ///  
  contrastvar(RS_diff)
```

- Not comparable due to different confounding (age) distribution.
- We can ask for a contrast between the at options.

# Internal Age-Standardisation

- Take average within each deprivation group
  - 25,618 in least deprived and 9672 in most deprived.

$$\bar{R}(t|X=x, Z_{X=x}) = \frac{1}{N_{X=x}} \sum_{i=1}^{N_{X=x}} \hat{R}_i(t|X=x, Z_i)$$

```
standsurv, survival timevar(tt) ci    ///  
at1(., atif(dep5==1))                ///  
at2(., atif(dep5==0))                ///  
atvar(RS_dep5 RS_dep1)               ///  
contrast(difference)                 ///  
contrastvar(RS_diff)
```

- Not comparable due to different confounding (age) distribution.
- We can ask for a contrast between the at options.

# Internal Age-Standardisation

- Take average within each deprivation group
  - 25,618 in least deprived and 9672 in most deprived.

$$\bar{R}(t|X=x, Z_{X=x}) = \frac{1}{N_{X=x}} \sum_{i=1}^{N_{X=x}} \hat{R}_i(t|X=x, Z_i)$$

```
standsurv, survival timevar(tt) ci    ///  
  at1(., atif(dep5==1))              ///  
  at2(., atif(dep5==0))              ///  
  atvar(RS_dep5 RS_dep1)             ///  
  contrast(difference)                ///  
  contrastvar(RS_diff)
```

- Not comparable due to different confounding (age) distribution.
- We can ask for a contrast between the at options.

# Internal Age-Standardisation

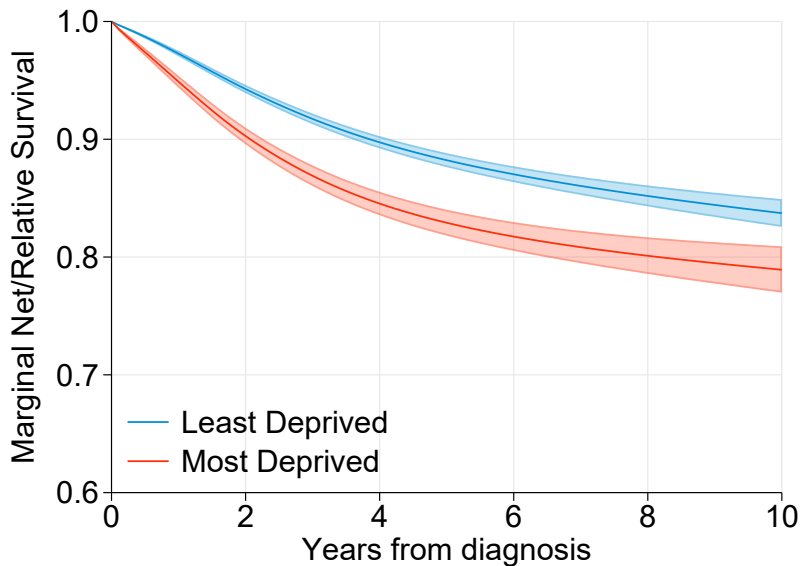
- Take average within each deprivation group
  - 25,618 in least deprived and 9672 in most deprived.

$$\bar{R}(t|X=x, Z_{X=x}) = \frac{1}{N_{X=x}} \sum_{i=1}^{N_{X=x}} \hat{R}_i(t|X=x, Z_i)$$

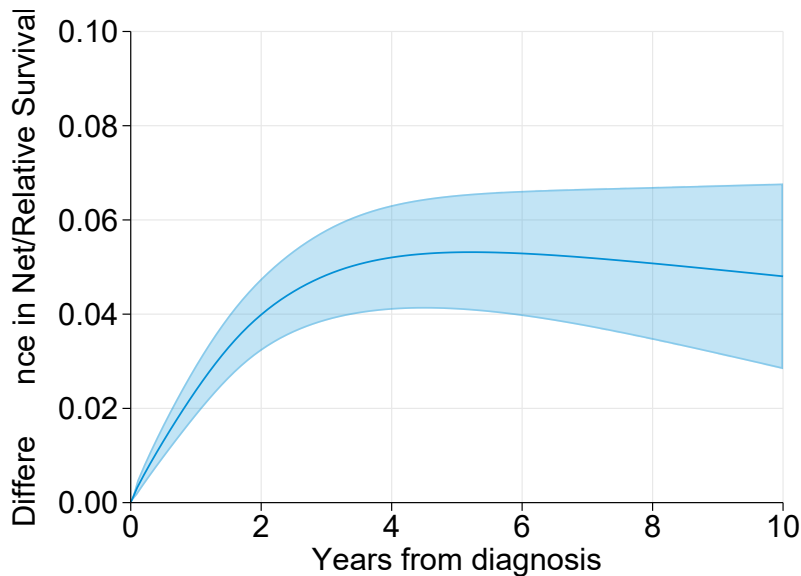
```
standsurv, survival timevar(tt) ci    ///  
  at1(., atif(dep5==1))              ///  
  at2(., atif(dep5==0))              ///  
  atvar(RS_dep5 RS_dep1)             ///  
  contrast(difference)                ///  
  contrastvar(RS_diff)
```

- Not comparable due to different confounding (age) distribution.
- We can ask for a contrast between the at options.

# Internally standardized (separately)



# Internally standardized (separately)





# Age-Standardisation (combined age distribution)

$$\bar{R}(t|X = x, Z) = \frac{1}{N} \sum_{i=1}^N \hat{R}_i(t|X = x, Z_i)$$

```
standsurv, survival timevar(tt) ci verbose          ///  
  atvar(RS_dep5_stand RS_dep1_stand)              ///  
  at1(dep5 1 agercs1_dep5=agercs1 agercs2_dep5=agercs2 agercs3_dep5=agercs3) ///  
  at2(dep5 0 agercs1_dep5 0 agercs2_dep5 0 agercs3_dep5 0)      ///  
  contrast(difference)                                         ///  
  contrastvar(RS_stand_diff)
```

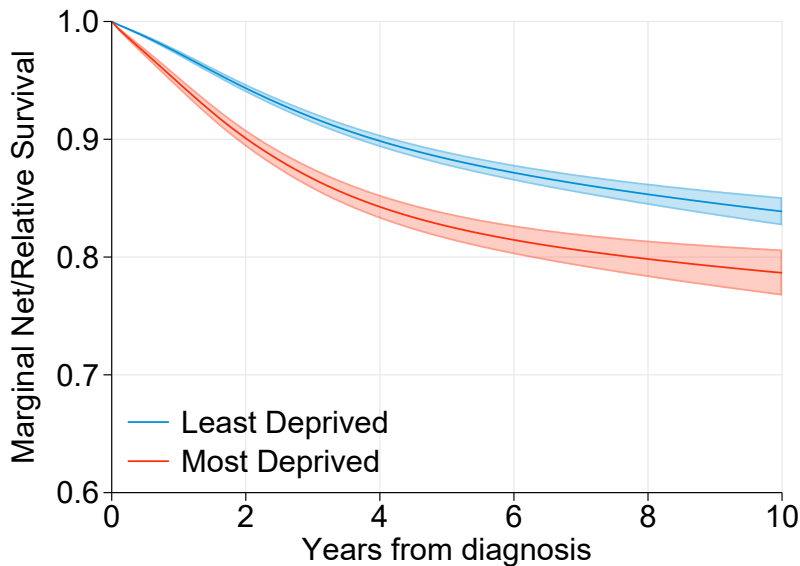
# Age-Standardisation (combined age distribution)

$$\bar{R}(t|X = x, Z) = \frac{1}{N} \sum_{i=1}^N \hat{R}_i(t|X = x, Z_i)$$

```
standsurv, survival timevar(tt) ci verbose          ///  
  atvar(RS_dep5_stand RS_dep1_stand)              ///  
  at1(dep5 1 agerics1_dep5=agerics1 agerics2_dep5=agerics2 agerics3_dep5=agerics3) ///  
  at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0)      ///  
  contrast(difference)                                          ///  
  contrastvar(RS_stand_diff)
```

- Comparable as age distribution is the same between the two groups.

# Age standardized (combined age distribution)



# External age-standardisation (ICSS)

- We often use the ICSS weights so that estimates are comparable between different studies.

$$\bar{R}(t|X=x, Z) = \frac{1}{N} \sum_{i=1}^N w_i \hat{R}_i(t|X=x, Z_i)$$

```
standsurv, survival timevar(tt) ci verbose    ///  
indweights(w)                               ///  
atvar(FR_dep5_extstand FR_dep1_extstand)    ///  
at1(., atif(dep5==1))                       ///  
at2(., atif(dep5==0))                       ///  
contrast(difference)                        ///  
contrastvar(RS_extstand_diff)
```

- Could fit separate models and still have comparable estimates.

# External age-standardisation (ICSS)

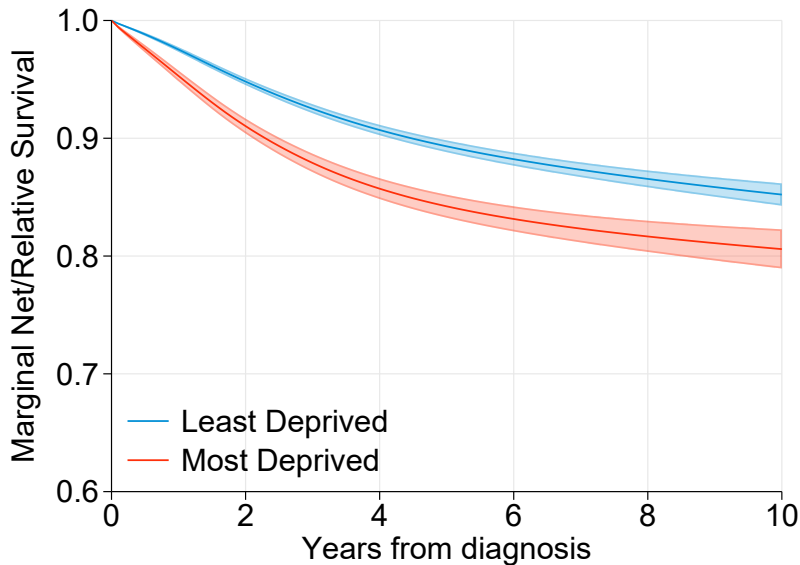
- We often use the ICSS weights so that estimates are comparable between different studies.

$$\bar{R}(t|X=x, Z) = \frac{1}{N} \sum_{i=1}^N w_i \hat{R}_i(t|X=x, Z_i)$$

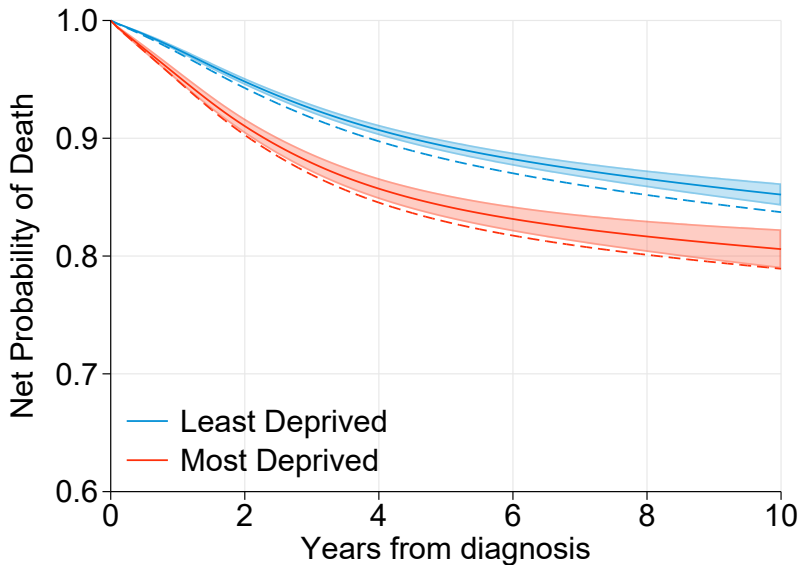
```
standsurv, survival timevar(tt) ci verbose    ///  
indweights(w)                               ///  
atvar(FR_dep5_extstand FR_dep1_extstand)    ///  
at1(., atif(dep5==1))                       ///  
at2(., atif(dep5==0))                       ///  
contrast(difference)                        ///  
contrastvar(RS_extstand_diff)
```

- Could fit separate models and still have comparable estimates.

# Externally Age standardized (ICSS)



# Externally Age standardized (ICSS)



# All-cause survival

- We model use relative/net survival to make fair comparisons between groups, by removing differences in other cause mortality rates.
- However, also useful to be able to quantify real world risks.
- This can be through all-cause survival or crude probabilities of death.
- We reuse the population mortality rates to predict all-cause survival.

$$\bar{S}(t|X = x, Z) = \frac{1}{N} \sum_{i=1}^N \hat{S}_i^*(t|Z_i, X = x) \hat{R}_i(t|Z_i, X = x)$$



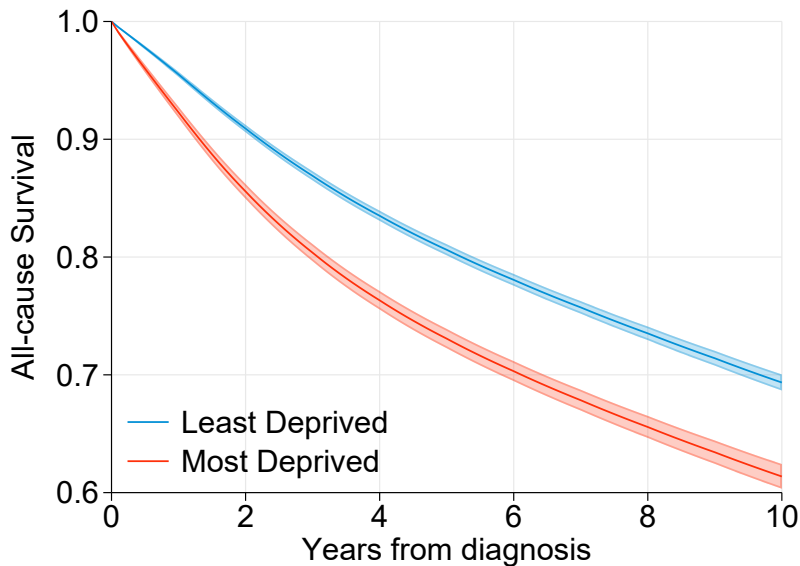
# Incorporate expected mortality rates

```
standsurv, survival timevar(tt) ci ///
  atvar(S_dep5 S_dep1) ///
  at1(dep5 1 agercs1_dep5=agercs1 agercs2_dep5=agercs2 agercs3_dep5=agercs3) ///
  at2(dep5 0 agercs1_dep5 0 agercs2_dep5 0 agercs3_dep5 0) ///
  expsurv(using(popmort_uk_regions_2017.dta) ///
    datediag(dx) ///
    ageddiag(agediag) ///
    pmrate(rate) ///
    pmage(age) ///
    pmyear(year) ///
    pmother(sex dep region) ///
    pmmaxyear(2016) ///
    at1(dep 5) ///
    at2(dep 1) ///
  ) ///
  contrast(difference) ///
  contrastvar(S_diff)
```

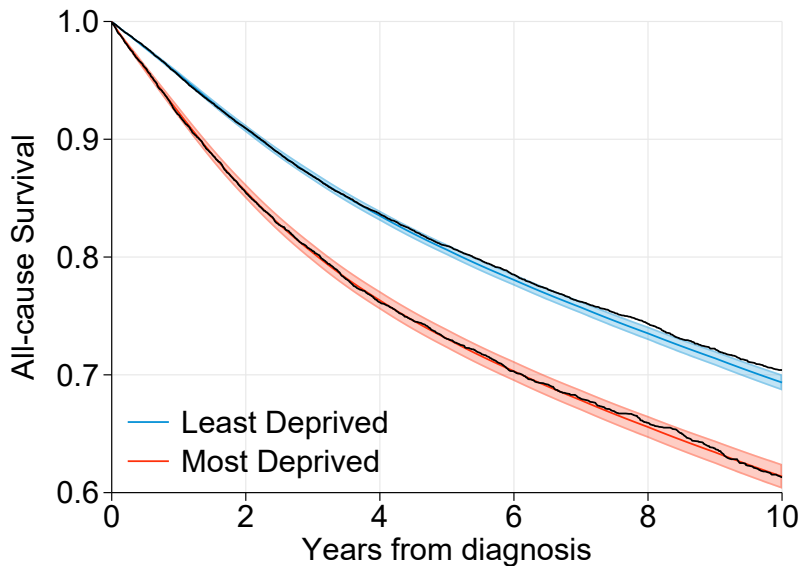
# Incorporate expected mortality rates

```
standsurv, survival timevar(tt) ci                                     ///  
  atvar(S_dep5 S_dep1)                                              ///  
  at1(dep5 1 agerics1_dep5=agerics1 agerics2_dep5=agerics2 agerics3_dep5=agerics3) ///  
  at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0)      ///  
  expsurv(using(popmort_uk_regions_2017.dta)                       ///  
    datediag(dx)                                                    ///  
    ageddiag(agediag)                                              ///  
    pmrate(rate)                                                    ///  
    pmage(age)                                                      ///  
    pmyear(year)                                                    ///  
    pmother(sex dep region)                                         ///  
    pmmaxyear(2016)                                                 ///  
    at1(dep 5)                                                       ///  
    at2(dep 1)                                                       ///  
  )                                                                    ///  
  contrast(difference)                                              ///  
  contrastvar(S_diff)                                              ///  
)
```

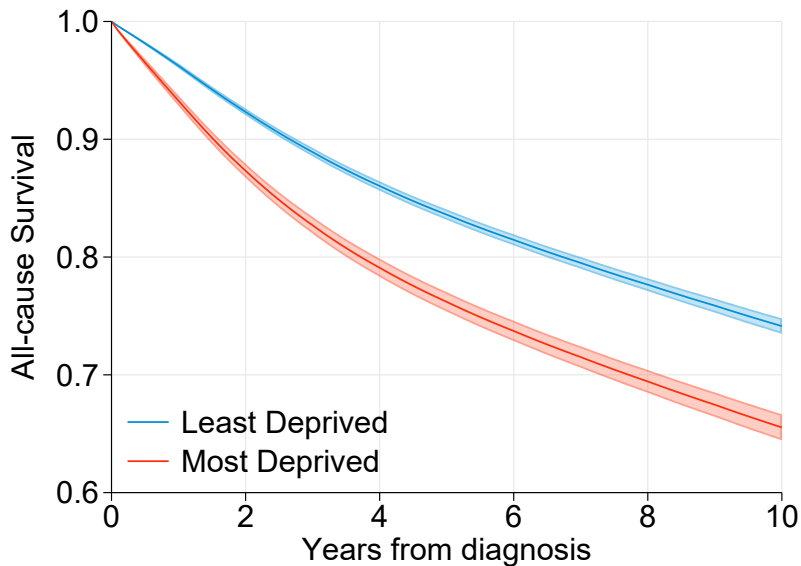
# Internal age-standardization



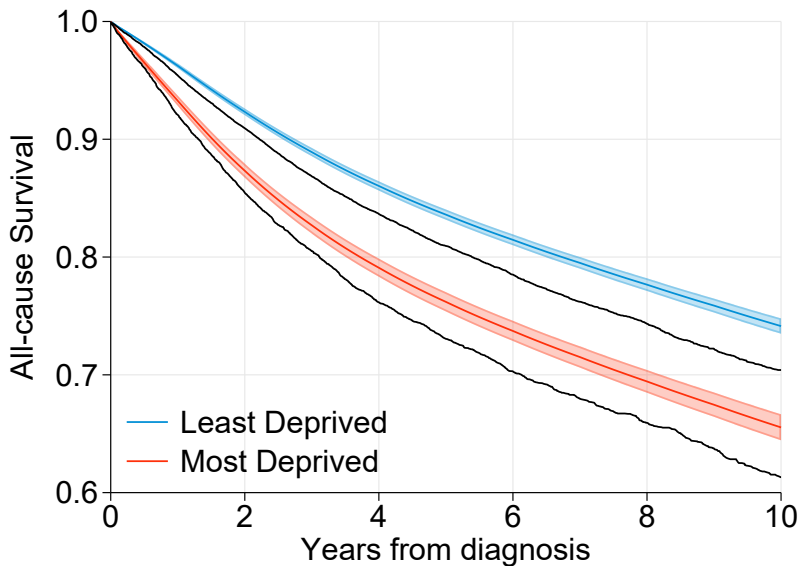
# Internal age-standardization



# Age-standardization (ICSS)



# Age-standardization (ICSS)



# Some comments

- The differences between the groups on the previous slides could be due to a combination of,
  - Differences in cancer mortality rates.
  - Differences in other cause mortality rates.
- This may be interest, but often we want isolate differences in cancer mortality rates.
- Will come back to this in second talk.

# Avoidable Deaths

- One way of trying to quantify differences between groups is to estimate avoidable deaths.

How many deaths would be avoided if we could reduce the excess mortality rate in one group to that observed in another group?

- Need to define what population this refers to.
- E.g. (average) number of individuals diagnosed in a calendar year,  $\tilde{N}$ .



# Avoidable Deaths 2

## Avoidable Deaths

$$\tilde{N} \left[ \left( 1 - E \left[ S^*(t|X, Z) R(t|X, Z) \right] \right) - \left( 1 - E \left[ S^*(t|X, Z) R(t|X = x, Z) \right] \right) \right]$$

- This is the difference in the expected number of deaths and the expected number of deaths if everyone was unexposed.
- Note only manipulating  $X$  for relative survival and not expected survival.

# Avoidable Deaths

```
count if yeardiag>=2008
local N_AD = 'r(N)'/5

standsurv if yeardiag>=2008, failure timevar(tt) ci          ///
  atvar(deaths_total deaths_dep10)                          ///
  at1(.)                                                      ///
  at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0) ///
  expsurv(using(popmort_uk_regions_2017.dta)                ///
    datediag(dx)                                              ///
    ageddiag(agediag)                                         ///
    pmrate(rate)                                              ///
    pmage(age)                                                ///
    pmyear(year)                                              ///
    pmother(sex dep region)                                   ///
    pmmaxyear(2016)                                           ///
  )                                                           ///
per('N_AD')                                                  ///
atref(2)                                                      ///
contrast(difference)                                          ///
contrastvar(AD1)
```

- Can also calculate avoidable deaths in age groups and sum.

# Avoidable Deaths

```
count if yeardiag>=2008
local N_AD = 'r(N)'/5

standsurv if yeardiag>=2008, failure timevar(tt) ci          ///
    atvar(deaths_total deaths_dep10)                        ///
    at1(.)                                                    ///
    at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0) ///
    expsurv(using(popmort_uk_regions_2017.dta)              ///
        datediag(dx)                                       ///
        ageddiag(agediag)                                  ///
        pmrate(rate)                                       ///
        pmage(age)                                         ///
        pmyear(year)                                       ///
        pmother(sex dep region)                           ///
        pmmaxyear(2016)                                    ///
    )                                                        ///
per('N_AD')                                                 ///
atref(2)                                                      ///
contrast(difference)                                         ///
contrastvar(AD1)
```

- Can also calculate avoidable deaths in age groups and sum.

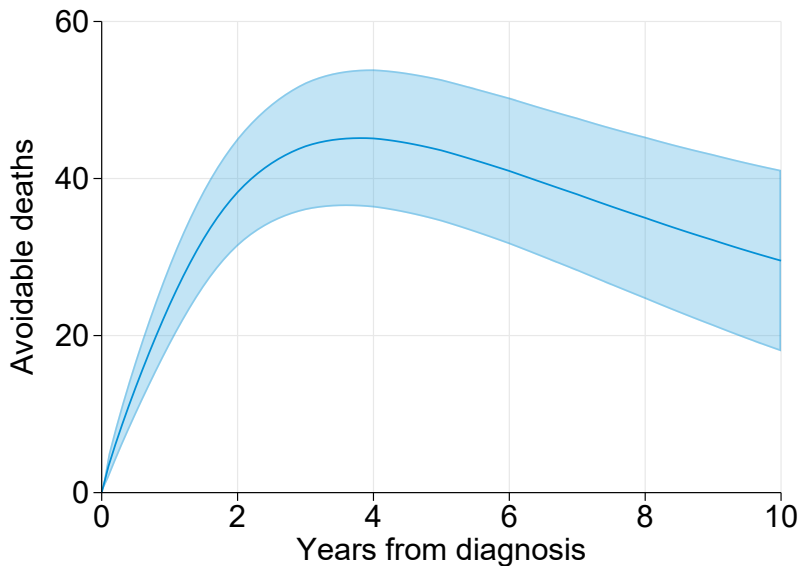
# Avoidable Deaths

```
count if yeardiag>=2008
local N_AD = 'r(N)'/5

standsurv if yeardiag>=2008, failure timevar(tt) ci          ///
    atvar(deaths_total deaths_dep10)                        ///
    at1(.)                                                    ///
    at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0) ///
    expsurv(using(popmort_uk_regions_2017.dta)              ///
        datediag(dx)                                         ///
        ageddiag(agediag)                                    ///
        pmrate(rate)                                         ///
        pmage(age)                                           ///
        pmyear(year)                                         ///
        pmother(sex dep region)                              ///
        pmmaxyear(2016)                                       ///
    )                                                         ///
per('N_AD')                                                  ///
atref(2)                                                      ///
contrast(difference)                                         ///
contrastvar(AD1)
```

- Can also calculate avoidable deaths in age groups and sum.

# Avoidable Deaths



# Attributable fraction (Relative Survival Framework)

- If we drop  $\tilde{N}$  from the avoidable deaths equation we have the numerator for the attributable fraction.

$$AF(t) = \frac{(1 - E[S^*(t|X, Z)R(t|X, Z)]) - (1 - E[S^*(t|X, Z)R(t|X = x, Z)])}{1 - E[S^*(t|X, Z)R(t|X, Z)]}$$

$$AF(t) = 1 - \frac{(1 - E[S^*(t|X, Z)R(t|X = x, Z)])}{1 - E[S^*(t|X, Z)R(t|X, Z)]}$$

- This is a non-linear function of different at() options.
- We use the `userfunction()` option.

# Attributable fraction code

```
mata:
function calcAF(at) {
  // at1() is F(t,Z): at2() is F(t|unexposed,Z)
  return(1 - at[2]/at[1])
}
end

standsurv, failure timevar(tt) ci verbose          ///
  atvar(Pdead Pdead_dep102)                        ///
  at1(.)                                            ///
  at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0) ///
  expsurv(using(popmort_uk_regions_2017.dta)      ///
    datediag(dx)                                  ///
    ageddiag(agediag)                             ///
    pmrate(rate)                                   ///
    pmage(age)                                     ///
    pmyear(year)                                   ///
    pmother(sex dep region)                       ///
    pmmaxyear(2016)                               ///
  )                                                ///
userfunction(calcAF)                             ///
userfunctionvar(AF)
```

# Attributable fraction code

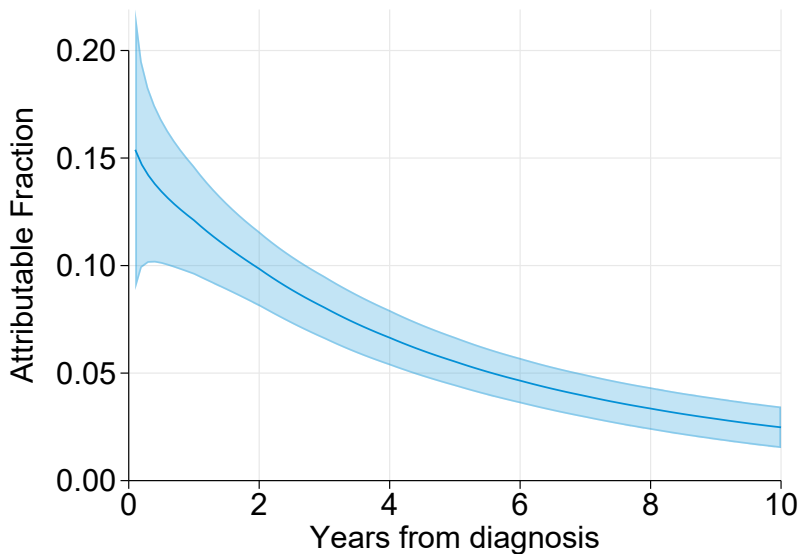
```
mata:  
function calcAF(at) {  
    // at1() is F(t,Z): at2() is F(t|unexposed,Z)  
    return(1 - at[2]/at[1])  
}  
end  
  
standsurv, failure timevar(tt) ci verbose          ///  
    atvar(Pdead Pdead_dep102)                      ///  
    at1(.)                                           ///  
    at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0) ///  
    expsurv(using(popmort_uk_regions_2017.dta)      ///  
        datediag(dx)                               ///  
        ageddiag(agediag)                          ///  
        pmrate(rate)                               ///  
        pmage(age)                                  ///  
        pmyear(year)                                ///  
        pmother(sex dep region)                    ///  
        pmmaxyear(2016)                            ///  
    )                                               ///  
    userfunction(calcAF)                           ///  
    userfunctionvar(AF)
```



# Attributable fraction code

```
mata:  
function calcAF(at) {  
  // at1() is F(t,Z): at2() is F(t|unexposed,Z)  
  return(1 - at[2]/at[1])  
}  
end  
  
standsurv, failure timevar(tt) ci verbose          ///  
  atvar(Pdead Pdead_dep102)                        ///  
  at1(.)                                             ///  
  at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0) ///  
  expsurv(using(popmort_uk_regions_2017.dta)      ///  
    datediag(dx)                                   ///  
    ageddiag(agediag)                             ///  
    pmrate(rate)                                   ///  
    pmage(age)                                     ///  
    pmyear(year)                                   ///  
    pmother(sex dep region)                       ///  
    pmmaxyear(2016)                               ///  
  )                                                 ///  
userfunction(calcAF)                               ///  
userfunctionvar(AF)
```

# Attributable Fraction



- I have calculated the percentage reduction in deaths among the whole population. Could restrict to the exposed using `if()/atif()` options.
- Could calculate AF using relative survival, but more awkward interpretation.

# Crude Probabilities

- Crude Probabilities enable us to partition the all-cause probability of death into deaths due to cancer and deaths due to other causes.
- This is in a relative survival framework. Those more familiar with competing risks will know this as the (cause-specific) cumulative incidence function.

$$F_c(t|X, Z) = E \left[ \int_0^t S^*(u|X, Z) R(u|X, Z) \lambda(u|X, Z) du \right]$$

$$F_o(t|X, Z) = E \left[ \int_0^t S^*(u|X, Z) R(u|X, Z) h^*(u|X, Z) du \right]$$

$$F_c(t|X, Z) + F_o(t|X, Z) = F(t|X, Z) = 1 - S^*(t|X, Z) R(t|X, Z)$$

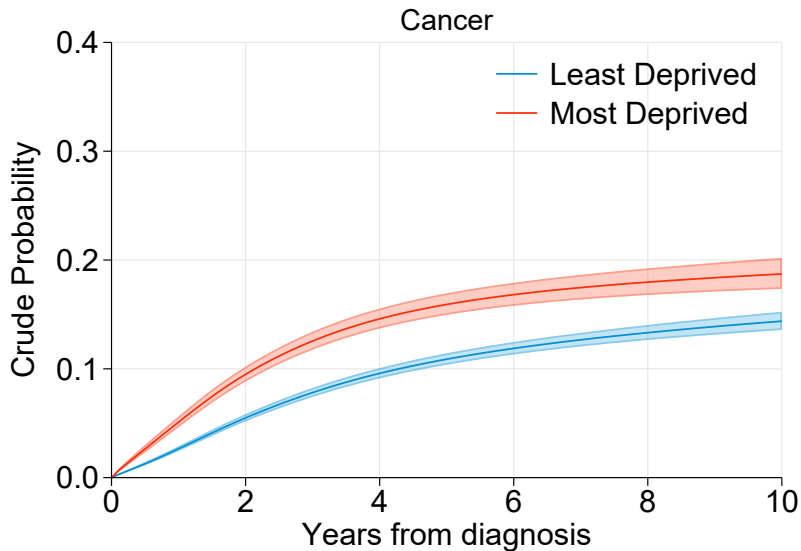
# Crude Probabilities code

```
standsurv, crudeprob timevar(tt) ci verbose ode          ///
  atvar(CP_dep5 CP_dep1)                                ///
  at1(dep5 1 agerics1_dep5=agerics1 agerics2_dep5=agerics2 agerics3_dep5=agerics3) ///
  at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0) ///
  expsurv(using(popmort_uk_regions_2017.dta)            ///
    datediag(dx)                                         ///
    ageddiag(agediag)                                   ///
    pmrate(rate)                                         ///
    pmage(age)                                           ///
    pmyear(year)                                         ///
    pmother(sex dep region)                             ///
    pmmaxyear(2016)                                     ///
    at1(dep 5)                                           ///
    at2(dep 1)                                           ///
  )                                                       ///
  contrast(difference)                                  ///
  contrastvar(CP_diff)
```

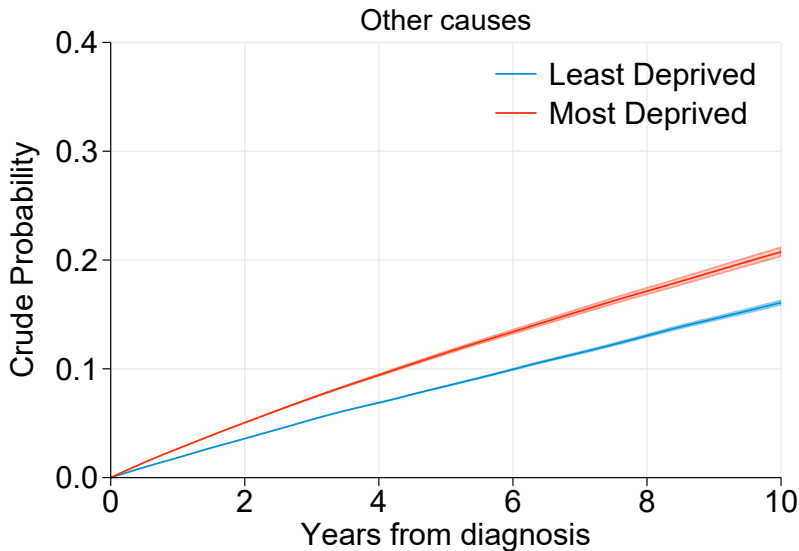
# Crude Probabilities code

```
standsurv, crudeprob timevar(tt) ci verbose ode ///
  atvar(CP_dep5 CP_dep1) ///
  at1(dep5 1 agerics1_dep5=agerics1 agerics2_dep5=agerics2 agerics3_dep5=agerics3) ///
  at2(dep5 0 agerics1_dep5 0 agerics2_dep5 0 agerics3_dep5 0) ///
  expsurv(using(popmort_uk_regions_2017.dta) ///
    datediag(dx) ///
    ageddiag(agediag) ///
    pmrate(rate) ///
    pmage(age) ///
    pmyear(year) ///
    pmother(sex dep region) ///
    pmmaxyear(2016) ///
    at1(dep 5) ///
    at2(dep 1) ///
  ) ///
  contrast(difference) ///
  contrastvar(CP_diff)
```

# Crude Probability Graphs

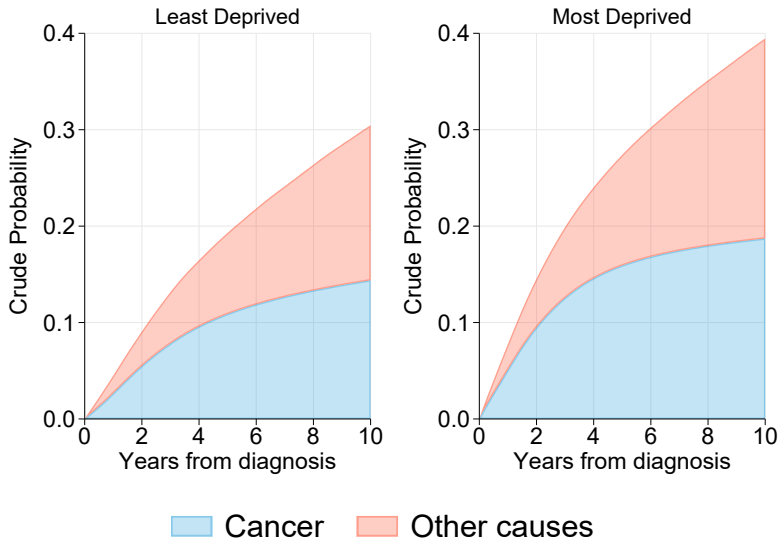


# Crude Probability Graphs





# Crude Probability Stacked Graphs



# Attributable fraction in crude probability framework

- Can estimate proportional reduction in cancer deaths.
- Can use cancer deaths as denominator...

$$\frac{E[F_c(t|X, Z)] - E[F_c(t|X = 0, Z)]}{E[F_c(t|X, Z)]}$$

- .... or all deaths

$$\frac{E[F_c(t|X, Z)] - E[F_c(t|X = 0, Z)]}{E[F_c(t|X, Z)] + E[F_o(t|X, Z)]}$$

- Just need to apply the `userfunction()` option again.

```
mata:  
function calcAF_cancer1(at) {  
    return((at[1] - at[3])/at[1])  
}  
end
```

```
mata:  
function calcAF_cancer2(at) {  
    return((at[1] - at[3])/(at[1] + at[2]))  
}  
end
```

# Life Expectancy

- We have previously presented work on the loss in expectation of life associated with a cancer diagnosis.
- This has generally been presented conditional estimates, which is useful as life expectancy as it is so strongly associated with age.
- Calculation of life-expectancy requires extrapolation of all-cause survival, but we can do this pretty well by combining extrapolations of expected and relative survival.
- Alternatively a restricted mean can be calculated. This is simply the area under the marginal survival curve.

$$RMST = E \left[ \int_0^{t^*} S^*(u|X, Z) R(u|X, Z) du \right]$$

- If  $t^*$  is large enough so the all-cause survival is (effectively) zero, this is an estimate of marginal life expectancy.

# Loss in Expectation of Life

- The marginal loss in expectation in life is the difference between the marginal life expectancy in the general population and the marginal life expectancy of the cancer population.

$$LEL = E \left[ \int_0^{t^*} S^*(u|X, Z) du \right] - E \left[ \int_0^{t^*} S^*(u|X, Z) R(u|X, Z) du \right]$$

- Can then manipulate exposures.

$$\left( E \left[ \int_0^{t^*} S^*(u|X=1, Z) du \right] - E \left[ \int_0^{t^*} S^*(u|X=1, Z) R(u|X=1, Z) du \right] \right) - \left( E \left[ \int_0^{t^*} S^*(u|X=0, Z) du \right] - E \left[ \int_0^{t^*} S^*(u|X=0, Z) R(u|X=0, Z) du \right] \right)$$

# Code

```
gen tstar = 100 in 1
standsurv, rmst timevar(tstar) ci                                     ///
  atvar(LifeExp5 LifeExp1)                                           ///
  at1(dep5 1 agercs1_dep5=agercs1 agercs2_dep5=agercs2 agercs3_dep5=agercs3) ///
  at2(dep5 0 agercs1_dep5 0 agercs2_dep5 0 agercs3_dep5 0)         ///
  expsurv(using(popmort_uk_regions_2017.dta)                       ///
    datediag(dx)                                                    ///
    ageddiag(agediag)                                              ///
    pmrate(rate)                                                    ///
    pmage(age)                                                      ///
    pmyear(year)                                                    ///
    pmother(sex dep region)                                         ///
    pmmaxyear(2016)                                                 ///
    at1(dep 5)                                                       ///
    at2(dep 1)                                                       ///
    expsurvvars(PopExp5 PopExp1)                                    ///
  )                                                                    ///
  contrast(difference)                                              ///
  contrastvar(LifeExp_diff)                                         ///

gen LEL1 = PopExp1 - LifeExp1
gen LEL5 = PopExp5 - LifeExp5
```

# Code

```
gen tstar = 100 in 1
standsurv, rmst timevar(tstar) ci                                     ///
    atvar(LifeExp5 LifeExp1)                                          ///
    at1(dep5 1 agercs1_dep5=agercs1 agercs2_dep5=agercs2 agercs3_dep5=agercs3) ///
    at2(dep5 0 agercs1_dep5 0 agercs2_dep5 0 agercs3_dep5 0)        ///
    expsurv(using(popmort_uk_regions_2017.dta)                      ///
        datediag(dx)                                                ///
        ageddiag(ageddiag)                                          ///
        pmrate(rate)                                                ///
        pmage(age)                                                  ///
        pmyear(year)                                                ///
        pmother(sex dep region)                                     ///
        pmmaxyear(2016)                                             ///
        at1(dep 5)                                                  ///
        at2(dep 1)                                                  ///
        expsurvvars(PopExp5 PopExp1)                                ///
    )                                                                ///
    contrast(difference)                                           ///
    contrastvar(LifeExp_diff)                                       ///

gen LEL1 = PopExp1 - LifeExp1
gen LEL5 = PopExp5 - LifeExp5
```

# Code

```
gen tstar = 100 in 1
standsurv, rmst timevar(tstar) ci                                     ///
    atvar(LifeExp5 LifeExp1)                                         ///
    at1(dep5 1 agercs1_dep5=agercs1 agercs2_dep5=agercs2 agercs3_dep5=agercs3) ///
    at2(dep5 0 agercs1_dep5 0 agercs2_dep5 0 agercs3_dep5 0)       ///
    expsurv(using(popmort_uk_regions_2017.dta)                     ///
        datediag(dx)                                               ///
        ageddiag(agediag)                                         ///
        pmrate(rate)                                               ///
        pmage(age)                                                 ///
        pmyear(year)                                               ///
        pmother(sex dep region)                                    ///
        pmmaxyear(2016)                                            ///
        at1(dep 5)                                                 ///
        at2(dep 1)                                                 ///
        expsurvvvars(PopExp5 PopExp1)                             ///
    )                                                                ///
    contrast(difference)                                           ///
    contrastvar(LifeExp_diff)                                       ///

gen LEL1 = PopExp1 - LifeExp1
gen LEL5 = PopExp5 - LifeExp5
```

# LEL estimates

```
list PopExp1 LifeExp1 LEL1 P_LEL1 in 1
```

	PopExp1	LifeExp1	LEL1	P_LEL1
1.	27.210407	23.42349	3.786917	.1391716

```
. list PopExp5 LifeExp5 LEL5 P_LEL5 in 1
```

	PopExp5	LifeExp5	LEL5	P_LEL5
1.	23.244415	19.456234	3.788181	.1629717



# Other Things

- `standsurv` works with standard survival models, relative survival models and competing risk models.
- Could have written this whole talk from more of a causal perspective using potential outcomes.
- Works with `streg`, `stpm2`, `strcs` models.

# Other Things 2

## Key points

- Ability to restrict standardization to subgroups.
- Ability to manipulate exposures.
- Ability to standardize to external population.
- Ability to add weights.
- Range of different marginal measures.
- Contrasts between marginal measures.
- Non-linear transformation of marginal effects.
- Standard errors calculated using delta methods or M-estimation.
- Centiles, marginal hazards, RMST.....

# Other Things 3

- Competing risks:
  - Can use different models for different causes. E.g. Weibull for one cause and flexible parametric model for another
  - models allows different time-scales (Nicks talk).
  - Standardized CIFs and other measures.
- Written in Mata, analytic derivatives, so fast.
- Different weights for different at() options enables mediation analysis (See Betty's talk)

More information on `standsurv` available at

<https://pclambert.net/software/standsurv/>