How to deal with confounders in survival curves; a short (useful) review
working progress

Rino Bellocco, Sc.D.

Department of Statistics and Quantitative Methods
University of Milano-Bicocca
&
Department of Medical Epidemiology and Biostatistics
Karolinska Institutet

September 17, 2026

The core distinction

Kaplan–Meier describes observed survival by group

Unadjusted Kaplan–Meier

Answers: “What survival did each group experience?”

```
sts graph, by(treatment)
sts test treatment
```

Adjusted survival estimate

Answers: “What would survival look like if groups had the same confounder distribution?”

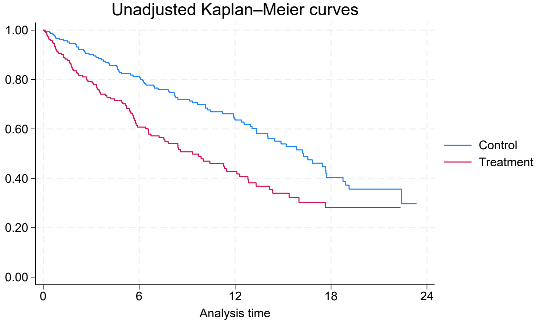
```
stcox treatment age sex stage
stcurve, survival at1(treatment=0)
         at2(treatment=1)
```

Background

- In clinical trials we often see a Kaplan-Meier curve plotted, simple to interpret (though there can be confusion with competing risks).
- In observational studies, we deal with confounding and take into account confounders in the model (Cox, parametric models).
- An adjusted hazard ratio often is estimated as long as we are ok with
  - the proportional hazards assumption
  - inclusion of all relevant confounders
  - modelling assumptions (non-linear effect, interactions etc)
- However, hazard ratios are more difficult to interpret and there are further problems when using hazard ratios as causal effects (Hernan 2010, Aaalen et al. 2015). Risks are much easier to interpret than rates and so quantifying the difference on the survival scale can be desirable

```
stset time, failure(death==1)
sts graph , by(treatment)
sts test treatment
stcox treatment
```

- I will explore differences between who received treatment and those who did not. This is our exposure, but as this is observational study we know that any association we see may be due to confounding. But.. it is always good to start with a Kaplan-Meier plot

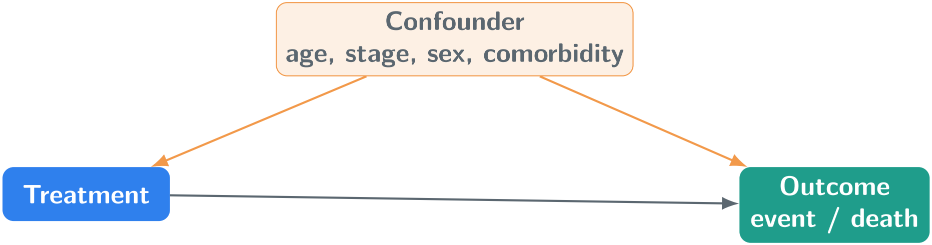


- This plot shows that those receiving treatment had worse survival. If we fit a proportional hazards model with hormon as the only covariate we get the following

```
                          (Std. err. adjusted for 400 clusters in id)
------------------------------------------------------------------------------
             |               Robust
          _t | Haz. ratio   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
             |
  Treatment  |   1.747152   .2633694     3.70   0.000     1.300225      2.3477
------------------------------------------------------------------------------
```

A confounder is a common cause

Adjustment targets a comparison that is less distorted by baseline imbalance



- The hazard ratio indicates that there is a 75% higher mortality rate in those receiving therapy.
- As this is an observational study we should not stop there and conclude that therapy is bad
- If treatment groups differ in age, sex, or disease stage, the raw KM gap might mix treatment effect with baseline risk, and therefore we do not have a fair comparison.

-

```
 Treatment |          Disease stage
     group |   Stage I   Stage II  Stage III |     Total
-----------+---------------------------------+----------
   Control |        85         74         54 |       213
           |     39.91      34.74      25.35 |    100.00
-----------+---------------------------------+----------
 Treatment |        51         52         84 |       187
           |     27.27      27.81      44.92 |    100.00
-----------+---------------------------------+----------
     Total |       136        126        138 |       400
           |     34.00      31.50      34.50 |    100.00
```

-

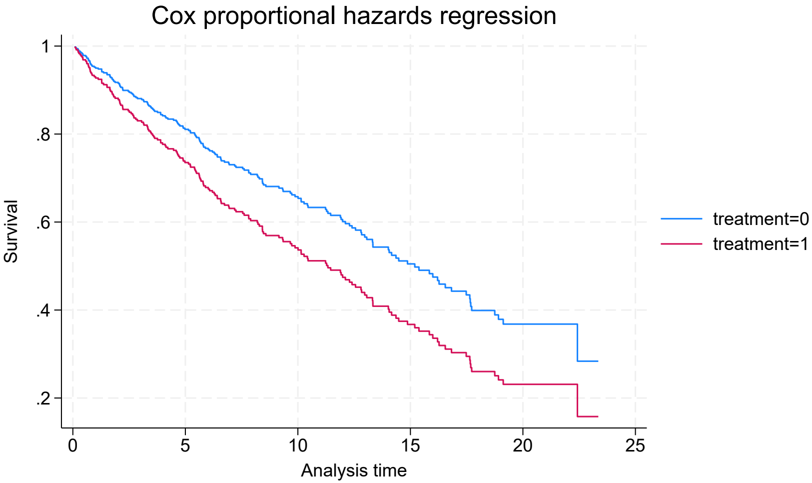
```
treatment |  Age Mean
----------+----------
  Control |  57.87793
Treatment |  61.26738
----------+----------
    Total |   59.4625
---------------------
```

```
stcox i.treatment, vce(robust)
```

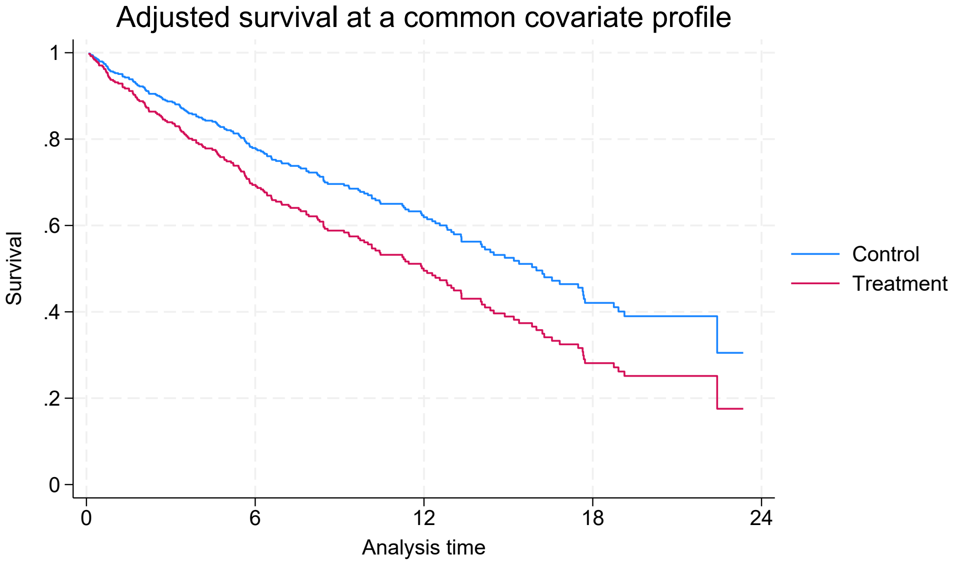
```
                               (Std. err. adjusted for 400 clusters in id)
------------------------------------------------------------------------------
             |               Robust
          _t | Haz. ratio   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
   treatment |
  Treatment  |   1.465324   .2307076     2.43   0.015     1.076258    1.995037
         age |   1.022646   .0084382     2.71   0.007     1.006241    1.039319
             |
         sex |
       Male  |    1.11067     .16993     0.69   0.493     .8229114    1.499053
             |
       stage |
   Stage II  |   1.257923   .2481985     1.16   0.245     .8544859    1.851838
  Stage III  |   1.693932    .317699     2.81   0.005     1.172876     2.44647
------------------------------------------------------------------------------
```

- Things have now a bit changed, the adjusted hazard ratio is now attenuated is (1.47) indicating a smaller detrimental effect
- I will try to understand what this hazard ratio means in terms of survival
- First I will replicate what stcurve does and then obtain the standardized survival functions

```
stcurve, survival at1(treatment=0) at2(treatment=1)
```



```
stcurve, survival ///
    at1(treatment=0 age=60 sex=0 stage=2) ///
    at2(treatment=1 age=60 sex=0 stage=2) ///
    xlabel(0(6)24) ///
    ylabel(0(.2)1) ///
    title("Adjusted survival at a common covariate profile") ///
    legend(order(1 "Control" 2 "Treatment"))
```


Adjusted survival at a common covariate profile
Survival
1
.8
.6
.4
.2
0
0
6
12
18
24
Analysis time
Control
Treatment

Standardized survival curves

- We want to calculate the expected survival under two counterfactuals. One where everyone has treatment and one where everybody does not take it (Sjolander 2016) for a nice description of these issues and implementation in an R package, stdreg.
- We also want contrasts of these standardized curves, for example a difference in standardized survival curves

$$E\left(S(t|X=1,Z\right)-E\left(S(t|X=0,Z\right)$$

Standardized survival curves

- We also want contrasts of these standardized curves, for example a difference in standardized survival curves

$$E\left(S(t|X=1,Z\right) - E\left(S(t|X=0,Z\right)$$

- X is the exposure of interest and Z's are the confounders. We are interested in the expectation over the distribution of Z with the key point being this distribution is forced to be the same for $X = 1$ and $X = 0$
- If our model is sufficent for confounding control then the above formula gives an average causal effect.
- To estimate the difference in the standardized curves we need to generate the two standardized survival curves. In each of these we predict as many survival curves as there are observations in the data set and then take the average of these curves. The only difference is that in one we make everyone be exposed (treatment=1) and in the other we make everybody be unexposed (treatment=0).

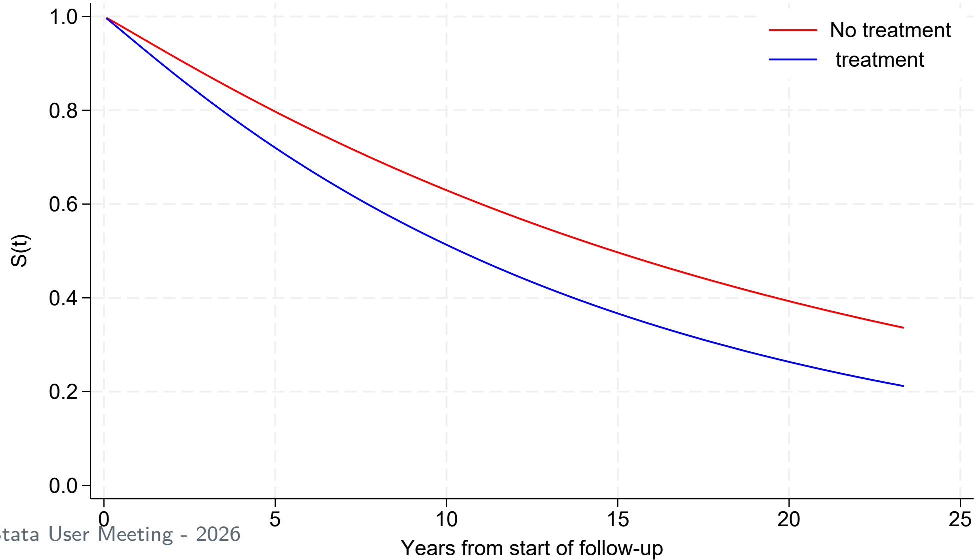
Standardized survival curves - ack P. Lambert

$$\frac{1}{N}\sum_{i=1}^{N} S(t|\text{treatment}{=}1, \text{age}_i, \text{sex}_i, \text{stage}_i, \text{pr\_1}_i) - \frac{1}{N}\sum_{i=1}^{N} S(t|\text{treatment}{=}0, \text{age}_i, \text{sex}_i, \text{stage}_i, \text{pr\_1}_i)$$

Standardized survival curves

```
streg treatment age sex stage, dist(weibull)
standsurv S0 S1, surv timevar(_t) ci                    ///
                  frame(f2, replace)                     ///
                  at1(treatment 0) at2(treatment 1)            ///
                  contrast(difference) contrastvar(Sdiff)
frame f2 {
 twoway (line S0 _t, sort lcolor(red))                                ///
        (line S1  _t, sort lcolor(blue))                              ///
         , legend(order(1 "No treatment" 2 " treatment")              ///
         ring(0) cols(1) pos(1))                                      ///
          ylabel(0 (0.2)1,angle(h) format(%3.1f))                    ///
          ytitle("S(t)") ///
          xtitle("Years from start of follow-up")
 }
```

Standardized survival curves



References

- Aalen O.O., Cook R.J. Roysland K. Does Cox analysis of a randomized survival study yield a causal treatment effect? *Lifetime data analysis*, **2015;21:579-593**
- Hernan M.A. The hazards of hazard ratios. Epidemiology **2010;21:13-15**
- Sjolander A. Regression standardization with the R package stdReg. *European Journal of Epidemiology*, **2016;31:563-574**