coxph

yimin chen

2023-12-07

In these data a subject changes exposure status from not bereaved to bereaved when his or her spouse dies. The first stage of the analysis therefore is to partition each follow–up into a record describing the period of follow-up pre–bereavement and (for subjects who were bereaved during the study) the period post–bereavement.

```r
## Creating times relativ to spouse death (year=0)
brv2 <- mutate(brv,
               id=NULL,
               y_before_sp_dth =  as.numeric(doe -dosp) / 365.24,
               y_after_sp_dth = as.numeric(dox - dosp) / 365.24)

## Splitting at spouse death (year=0)
brvSplit <- survSplit(brv2, cut = 0, end="y_after_sp_dth", start="y_before_sp_dth", id="id",event="fail

## Calculating risk times
brvSplit <- mutate(brvSplit,
                   t_sp_at_risk =   y_after_sp_dth - y_before_sp_dth,
                   brv = ifelse(y_after_sp_dth > 0, 1, 0))
```

```r
brvSplit2 <- mutate(brvSplit,
                    sex = as.factor(sex),
                    brv = as.factor(brv))
## Translate time scale from years from spouse death to ages
brvSplit3 <- brvSplit2 %>%
    mutate(age_sp_dth =  as.numeric(dosp - dob) / 365.24, # Age at spouse death
           age_start = age_sp_dth + y_before_sp_dth,      # Age at start of timeband
           age_end = age_sp_dth + y_after_sp_dth)         # Age at end of timeband

age_cat <- seq(70,100,5) # Split at these ages
brvSplit4 <- survSplit(brvSplit3, cut=age_cat, start="age_start", end="age_end", event="fail", zero = 0)

brvSplit4 <- mutate(brvSplit4,
                    t_at_risk = age_end- age_start, # Creating new time at risk
                    age = cut(age_end, age_cat))   # Creating age band category

## Calculate crude rates
survRate(Surv(t_at_risk, fail) ~ age, data=brvSplit4)
```

```
##                      age       tstop event       rate       lower       upper
```

```
## age=(75,80]   (75,80]  703.612419    45 0.06395566 0.04664970 0.08557771
## age=(80,85]   (80,85] 1184.684043   123 0.10382515 0.08628885 0.12387811
## age=(85,90]   (85,90]  490.021356    95 0.19386910 0.15685168 0.23699492
## age=(90,95]   (90,95]   55.090352    12 0.21782399 0.11255283 0.38049467
## age=(95,100] (95,100]    2.299858     3 1.30442857 0.26900453 3.81209383
```

```r
summary(coxph(Surv(age_start, age_end, fail) ~ brv,
              data = brvSplit4))
```

```
## Call:
## coxph(formula = Surv(age_start, age_end, fail) ~ brv, data = brvSplit4)
##
##   n= 1036, number of events= 278
##
##         coef exp(coef) se(coef)      z Pr(>|z|)
## brv1 -0.2070    0.8131   0.1390 -1.488    0.137
##
##      exp(coef) exp(-coef) lower .95 upper .95
## brv1    0.8131       1.23    0.6191     1.068
##
## Concordance= 0.511  (se = 0.014 )
## Likelihood ratio test= 2.26  on 1 df,   p=0.1
## Wald test            = 2.22  on 1 df,   p=0.1
## Score (logrank) test = 2.22  on 1 df,   p=0.1
```

also model these data using Cox regression. Provided we use the attained age as the time scale and split the data to obtain separate observations for the bereaved and non-bereaved person-time the following command will estimate the effect of bereavement adjusted for attained age.

```r
summary(coxph(Surv(age_start, age_end, fail) ~ brv + sex,
              data = brvSplit4))
```

```
## Call:
## coxph(formula = Surv(age_start, age_end, fail) ~ brv + sex, data = brvSplit4)
##
##   n= 1036, number of events= 278
##
##          coef exp(coef) se(coef)      z Pr(>|z|)
## brv1 -0.07842   0.92458  0.14245 -0.551 0.581971
## sex2 -0.47291   0.62318  0.13075 -3.617 0.000298 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      exp(coef) exp(-coef) lower .95 upper .95
## brv1    0.9246      1.082    0.6993    1.2224
## sex2    0.6232      1.605    0.4823    0.8052
##
## Concordance= 0.56  (se = 0.018 )
## Likelihood ratio test= 15.85  on 2 df,   p=4e-04
## Wald test            = 15.21  on 2 df,   p=5e-04
```

```
## Score (logrank) test = 15.51  on 2 df,   p=4e-04
```

```
summary(coxph(Surv(age_start, age_end, fail) ~ brv,
              data = brvSplit4))
```

```
## Call:
## coxph(formula = Surv(age_start, age_end, fail) ~ brv, data = brvSplit4)
##
##   n= 1036, number of events= 278
##
##         coef exp(coef) se(coef)      z Pr(>|z|)
## brv1 -0.2070    0.8131   0.1390 -1.488    0.137
##
##      exp(coef) exp(-coef) lower .95 upper .95
## brv1    0.8131       1.23    0.6191     1.068
##
## Concordance= 0.511  (se = 0.014 )
## Likelihood ratio test= 2.26  on 1 df,   p=0.1
## Wald test            = 2.22  on 1 df,   p=0.1
## Score (logrank) test = 2.22  on 1 df,   p=0.1
```

Use the Cox model to estimate the effect of bereavement separately for males and females and compare the estimates to those obtained using Poisson regression.

```
cox_model=coxph(Surv(age_start, age_end, fail) ~ brv + sex+group+disab+health,data = brvSplit4)
sum.cox=summary(coxph(Surv(age_start, age_end, fail) ~ brv + sex+group+disab+health,data = brvSplit4))
cox.coeff = sum.cox$coefficients
kable(cox.coeff,digit = 6)
```

	coef	exp(coef)	se(coef)	z	Pr(>|z|)
brv1	0.003182	1.003187	0.143946	0.022106	0.982363
sex2	-0.520998	0.593927	0.132939	-3.919082	0.000089
group	0.026896	1.027261	0.084823	0.317085	0.751179
disab	0.254538	1.289866	0.057233	4.447410	0.000009
health	-0.460763	0.630802	0.101012	-4.561458	0.000005

```
# Build the table for test
lrt = tibble(sum.cox$logtest)
wald = tibble(sum.cox$waldtest)
logrank = tibble(sum.cox$sctest)
test_table1 = cbind(lrt, wald, logrank)
rownames(test_table1) = c("test_statistics", "df", "pvalue")
colnames(test_table1) = c("Likelihood Ratio", "Wald", "Logrank")
test_table1 = round(test_table1, digits = 6)
kable(test_table1)
```

	Likelihood Ratio	Wald	Logrank
test_statistics	64.49756	71.05	73.31503
df	5.00000	5.00	5.00000

	Likelihood Ratio	Wald	Logrank
pvalue	0.00000	0.00	0.00000

```r
model.aic <- step(cox_model, direction = "both", k = 2)
```

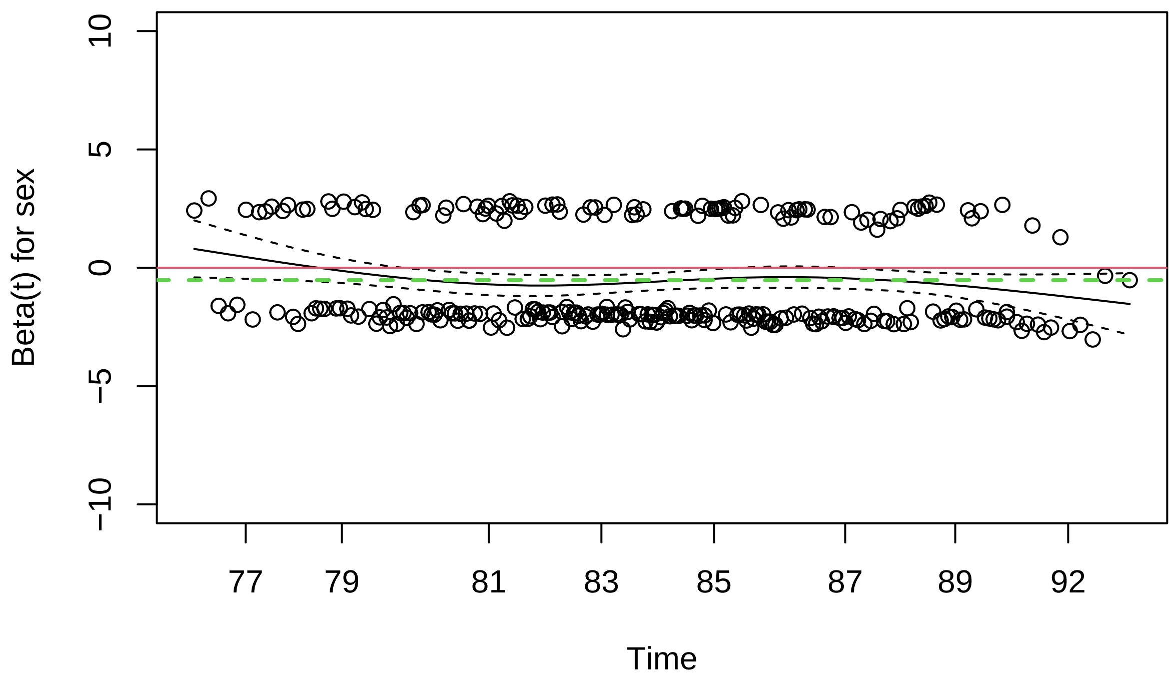
```
## Start:  AIC=2705.91
## Surv(age_start, age_end, fail) ~ brv + sex + group + disab +
##     health
##
##          Df    AIC
## - brv     1 2703.9
## - group   1 2704.0
## <none>      2705.9
## - sex     1 2719.9
## - disab   1 2721.7
## - health  1 2723.5
##
## Step:  AIC=2703.91
## Surv(age_start, age_end, fail) ~ sex + group + disab + health
##
##          Df    AIC
## - group   1 2702.0
## <none>      2703.9
## + brv     1 2705.9
## - sex     1 2718.9
## - disab   1 2719.8
## - health  1 2721.7
##
## Step:  AIC=2702.01
## Surv(age_start, age_end, fail) ~ sex + disab + health
##
##          Df    AIC
## <none>      2702.0
## + group   1 2703.9
## + brv     1 2704.0
## - sex     1 2717.7
## - disab   1 2717.9
## - health  1 2719.7
```

lowest aic: Surv(age_start, age_end, fail) ~ sex + disab + health

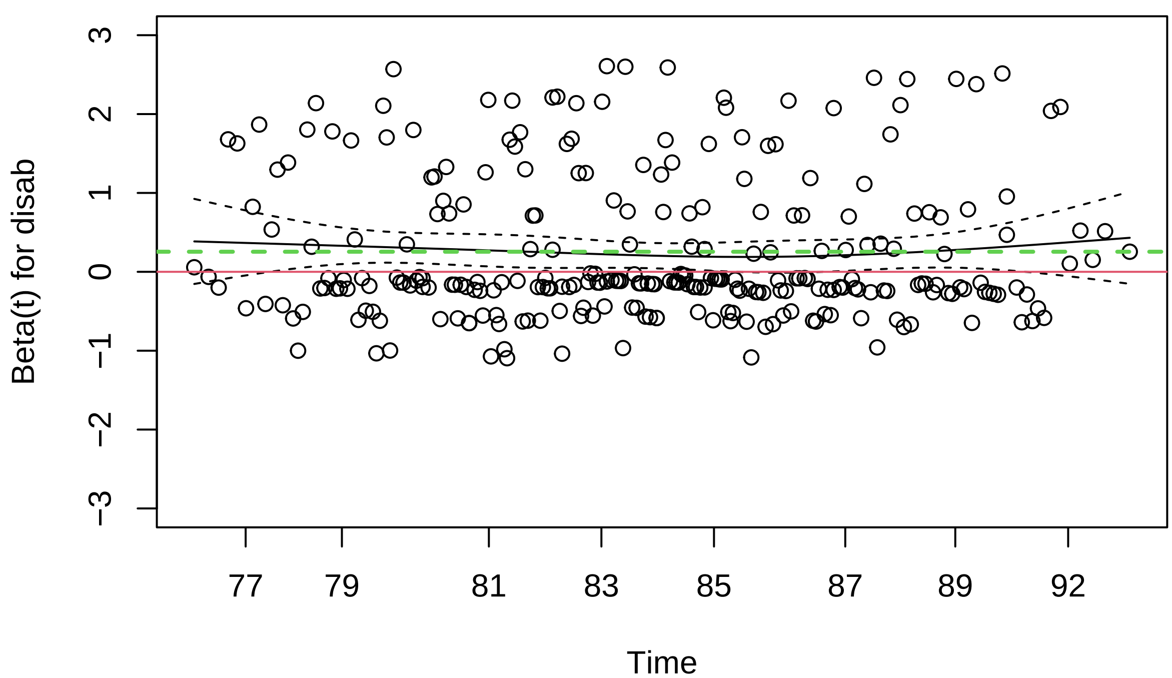
```r
sr_fit=coxph(Surv(age_start, age_end, fail) ~ sex + disab + health, data =brvSplit4)
czph <- cox.zph(coxph(Surv(age_start, age_end, fail) ~ sex + disab + health, data =brvSplit4))
czph
```

```
##        chisq df      p
## sex     3.16  1 0.0757
## disab   1.24  1 0.2652
## health 11.50  1 0.0007
## GLOBAL 13.11  3 0.0044
```

```
plot(czph[1],ylim=c(-10,10))
abline(a=0,b=0,col=2)
abline(h=sr_fit$coefficients[1],col=3,lwd=2,lty=2)
```




```
plot(czph[2],ylim=c(-3,3))
abline(a=0,b=0,col=2)
abline(h=sr_fit$coefficients[2],col=3,lwd=2,lty=2)
```

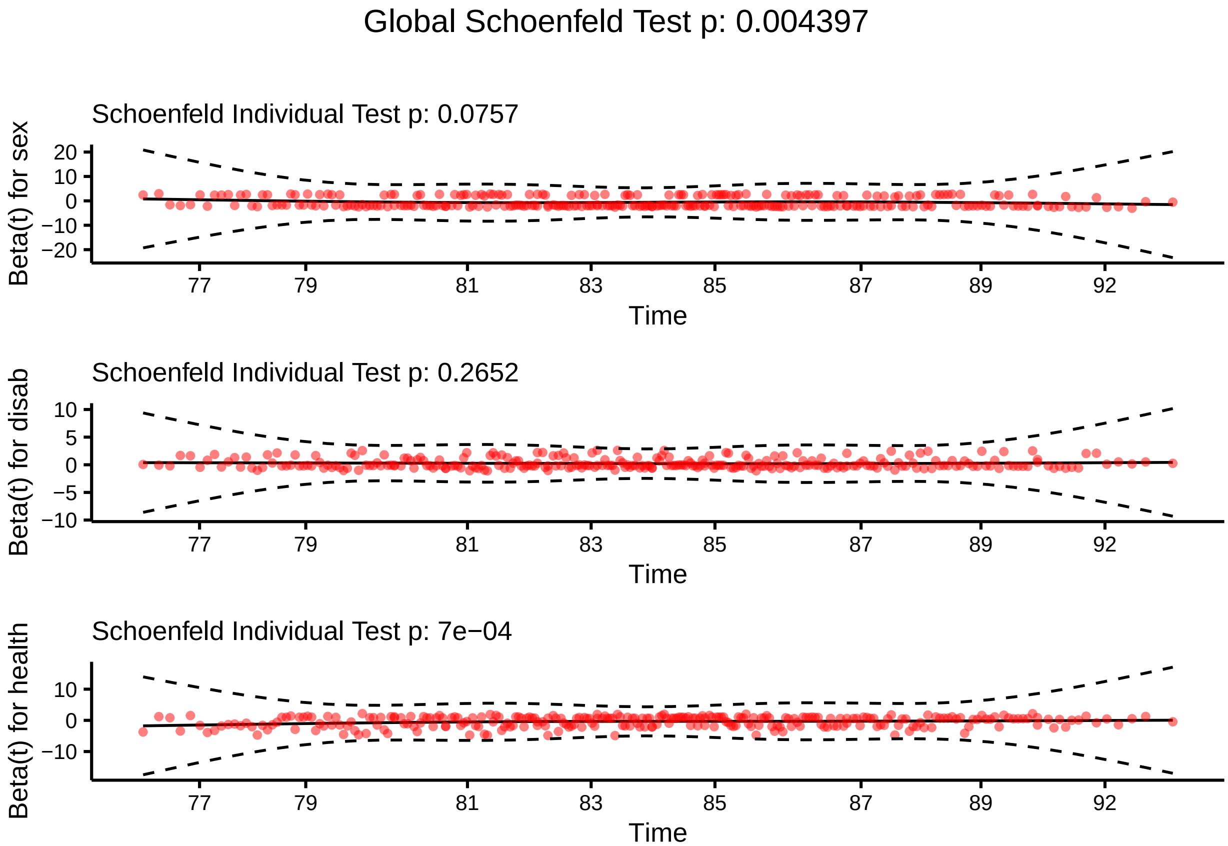



```
residual=ggcoxzph(czph,font.main=10,font.x=10,font.y=10, font.tickslab=8,point.alpha=0.5)
```

```
## Warning in regularize.values(x, y, ties, missing(ties), na.rm = na.rm):
```

```
## collapsing to unique 'x' values
```

```r
residual
```



The Global Schoenfeld Test has a p-value of 0.004397, which is below the common alpha level of 0.05, indicating that there is statistically significant evidence against the proportional hazards assumption across all covariates in the model. This suggests that the hazard ratios for the covariates in the model are not consistent over time.

The individual Schoenfeld Test for the 'sex' variable has a p-value of 0.0757, which is above the 0.05 threshold, suggesting that there is no significant evidence against the proportional hazards assumption for the 'sex' variable. The plot for 'sex' shows residuals scattered around the zero line without any apparent trend, which is in line with the proportional hazards assumption.

The individual Schoenfeld Test for 'disab' (presumably 'disability') has a p-value of 0.2652, also indicating no significant violation of the proportional hazards assumption for this variable.

Lastly, the individual Schoenfeld Test for 'health' has a very low p-value of 7e-04, indicating a significant violation of the proportional hazards assumption for this variable.

In summary, the Cox model may not be appropriate for the 'health' variable due to violation of the proportional hazards assumption. The 'sex' and 'disab' variables do not show evidence against this assumption. It may be necessary to explore time-varying covariates or alternative models for the 'health' variable to adequately model these data.

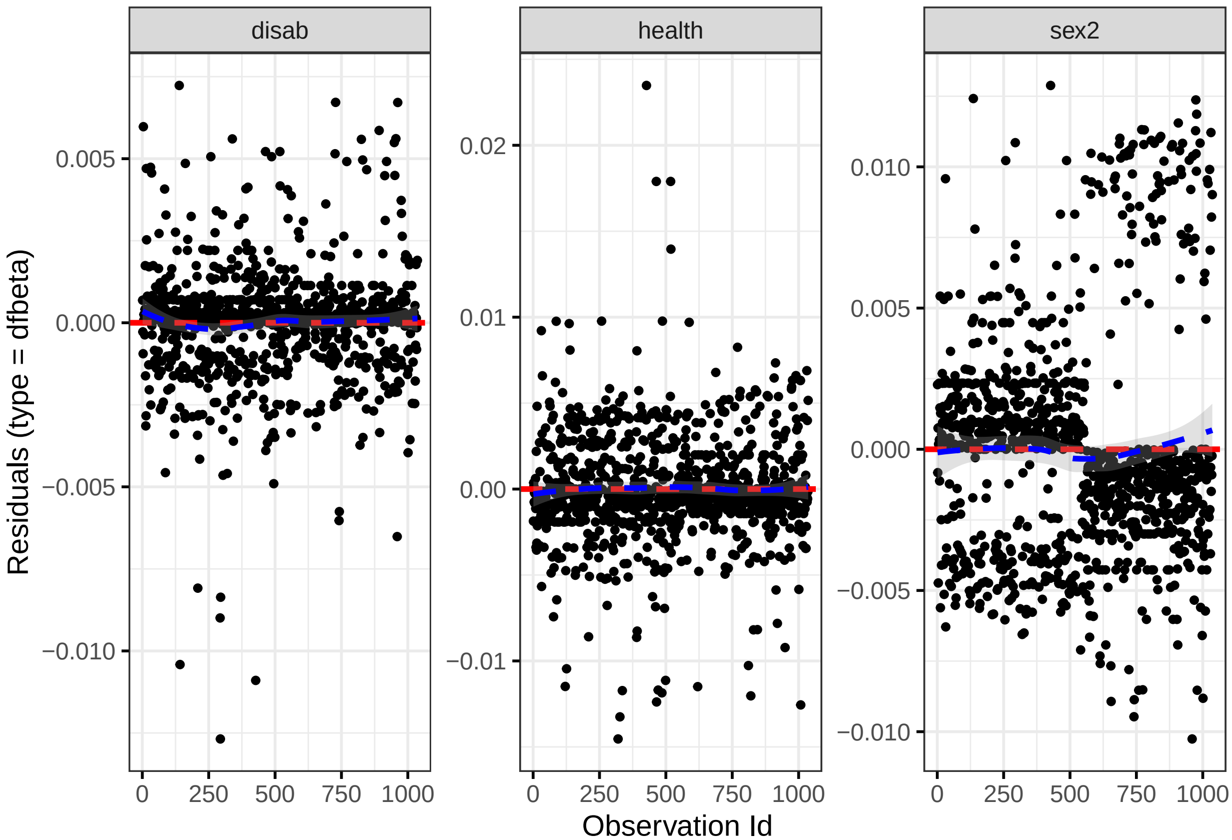
```r
outlier <- ggcoxdiagnostics(coxph(Surv(age_start, age_end, fail) ~ sex + disab + health, data =brvSplit
```

```
## Warning: `gather_()` was deprecated in tidyr 1.2.0.
## i Please use `gather()` instead.
## i The deprecated feature was likely used in the survminer package.
##   Please report the issue at <https://github.com/kassambara/survminer/issues>.
```

```
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
outlier
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



From the plots, it seems that for all covariates, there is no clear pattern of systematic influence, and no single observation appears to be particularly influential, as most of the dfbeta values are close to zero and within the confidence band.