Pitfalls of proportional hazards regression methods to assess sex-specific racial differences in survival in Eaton WW, Roth KB,..., Muñoz A. The Relationship of Mental and Behavioral Disorders to All-cause Mortality in a 27-year Follow-up of Four Epidemiologic Catchment Area Samples. Am J Epidemiol 2013

Alvaro Muñoz, PhD
Professor of Epidemiology

Survival after age 30 among Females by race

<table>
<thead>
<tr>
<th></th>
<th>Black</th>
<th></th>
<th>non-Black</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>deceased</td>
<td>alive</td>
<td>deceased</td>
<td>alive</td>
</tr>
<tr>
<td>N</td>
<td>841</td>
<td>1,805</td>
<td>2,950</td>
<td>3,358</td>
</tr>
<tr>
<td>Age at entry</td>
<td>[64, 72]</td>
<td>[30, 46]</td>
<td>[70, 77]</td>
<td>[30, 52]</td>
</tr>
<tr>
<td>Age at exit</td>
<td>[76, 85]</td>
<td>[52, 72]</td>
<td>[83, 90]</td>
<td>[53, 78]</td>
</tr>
<tr>
<td>P-Y on study</td>
<td>11,114</td>
<td>42,114</td>
<td>36,984</td>
<td>80,809</td>
</tr>
</tbody>
</table>
Survival after age 30 among Females by race

sts graph if Female==1, by(Black) xtitle(Years after Age 30) ytitle(Proportion Alive)

Kaplan-Meier survival estimates

Non-Black

Black

stcox Black if Female==1

* Poor summary: no difference by race and if anything, black race is protective
Cox as disguised Weibull

\[ T \sim \text{Weibull}(\alpha + \beta \cdot \text{black}, \sigma) \]

\[
\text{streg Black if Female==1, dist(weibull)}
\]

Weibull regression -- log relative-hazard form

No. of subjects = 8953
No. of failures = 3791
Time at risk = 171023.1643
Log likelihood = -909.33593
LR chi2(1) = 1.06
Prob > chi2 = 0.3042

|     | Haz. Ratio | Std. Err. | z    | P>|z| | 95% Conf. Interval |
|-----|------------|-----------|------|------|-------------------|
| Black | 0.9606234  | 0.0376984 | -1.02| 0.306| 0.889506 1.037427 |

\[
e^{-\beta/\sigma}
\]

Poor summary: inferences identical to Cox’s

Cox Regression allowing for time dependency

\[ h(t; \text{black}) = h(t; \text{non} - \text{black})e^{\beta + \gamma t} \]

\[
\text{stcox Black if Female==1, tvc(Black) texp( }_t \text{ ) nohr}
\]

No. of subjects = 8953
No. of failures = 3791
Time at risk = 171023.1643
Log likelihood = -28129.896
LR chi2(2) = 98.78
Prob > chi2 = 0.0000

|     | Coef. | Std. Err. | z    | P>|z| | 95% Conf. Interval |
|-----|-------|-----------|------|------|-------------------|
| main | \(\beta\) = |
| Black | 1.31299 | 0.1446371 | 9.08 | 0.000 | 1.029506 1.596473 |
| tvc  | \(\gamma\) = |
| Black | -0.0286116 | 0.0029663 | -9.65 | 0.000 | -0.034425 -0.0227978 |

Significant downward trend of HR
Hazard Ratio from a Cox Model allowing for time dependency

gen CoxHRBlacktoNonblackamongFemale= exp((main)Black + (tvc)Black * _t)
line CoxHRBlacktoNonblackamongFemale _t, sort xtitle(Years after Age 30)

Parametric Models: quantiles in addition to hazards
British MDs: Survival From Age 35, Continued

Doll et al., BMJ, 2004 doi:10.1136/bmj.38142.554479.AE

Fig 3 Survival from age 35 for continuing cigarette smokers and lifelong non-smokers among UK male doctors born 1900-1909, with percentages alive at each decade of age

Parametric Models

GG conventional \( T \sim GG(\alpha + \beta \; black, \sigma, \kappa) \)

times are proportional: \( \exp(\beta) \)

streg Black if Female==1, dist(gamma) time /*to display beta, sigma, kappa */

| _t | Coef. | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|----|-------|-----------|-------|------|---------------------|
| Beta Black | .0310971 | .0069531 | 4.47  | 0.000 | .0174692 -.044725 |
| alpha _cons | 4.119847 | .0062218 | 662.17 | 0.000 | 4.107652 -.044725 |
| /ln_sig | -1.914397 | .0348651 | -54.91 | 0.000 | -1.982731 -.846063 |
| /kappa | 2.499254 | .1083305 | 23.07 | 0.000 | 2.286933 2.711578 |
| sigma | .1474307 | .0051402 | 30.74 | 0.000 | .1376927 .1578575 |

exp(0.031) = 1.03

*For Non-black Females

Poor summary: Blacks live 3% longer

gen convGGSurvNonblackamongFemale= gammaptail(e(kappa)\(^{-2}\),e(kappa)\(^{-2}\))*\( \exp(\alpha \_cons) \)*ageexitm30)*\( \exp(\alpha \_cons) \)*ageexitm30)*e(kappa)/e(sigma))
replace convGGSurvNonblackamongFemale= 1-convGGSurvNonblackamongFemale if e(kappa)<0
Fit of conventional GG to non-parametric survival for Non-black Females

sts graph if Female==1 & Black==0, surv xtitle(Years after Age 30) ytitle(% Alive) ///
title(Nonblack Female) addplot(line convGGSurvNonblackamongFemale ageexitm30, sort)

Nonblack Female

Conventional GG
Non-parametric

Poor fit for Non-blacks

Fit of conventional GG to non-parametric survival for Black Females

sts graph if Female==1 & Black==1, surv xtitle(Years after Age 30) ytitle(% Alive) ///
title(Black Female) addplot(line convGGSurvBlackamongFemale ageexitm30, sort)

Black Female

Non-parametric
Conventional GG

and Poor fit for Blacks too
GG non-conventional (saturated)

\[ T \sim GG(\alpha + \beta \text{black}, \exp(\gamma_0 + \gamma_1 \text{black}), \kappa_0 + \kappa_1 \text{black}) \]

streg Black if Female==1, anc(Black) anc2(Black) dist(gamma)

|       | _t | Coef.   | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|-------|----|---------|-----------|-------|--------|---------------------|
| _t    |    | 0.0786024 | 0.0242907 | 3.24  | 0.001  | 0.0309936 - 0.1262113 |
| \alpha| cons| 4.107777 | 0.006334  | 648.53| 0.000   | 4.095363 - 4.120192  |

|       | ln_sig | _t    | Coef.   | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|-------|--------|-------|---------|-----------|-------|--------|---------------------|
| \gamma | Black  | -0.0521234 | 0.1767268 | -0.29  | 0.768  | -0.3985016 - 0.2942549 |
| \gamma | cons   | -1.918588 | 0.0342164 | -56.07 | 0.000   | -1.985651 - -1.851525 |

|       | kappa | _t    | Coef.   | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|-------|-------|-------|---------|-----------|-------|--------|---------------------|
| \kappa | Black | 1.266314 | 0.6381863 | 1.98   | 0.047  | 0.0154921 - 2.517136 |
| \kappa | cons  | 2.111767 | 0.10096  | 20.92  | 0.000   | 1.913889 - 2.309645  |

*For Non-black Females

gen convGGSurvNonblackamongFemale=
gammaptail(e(kappa)^(-2), (e(kappa)^(-2))*exp(_t_cons)*(ageexitm30)^e(kappa)/e(sigma))
replace convGGSurvNonblackamongFemale= 1-convGGSurvNonblackamongFemale if e(kappa)<0

---

Fit of non-conventional GG to non-parametric survival for Non-black Females

sts graph if Female==1 & Black==0, surv xtitle(Years after Age 30) ytitle(% Alive) ///
title(Nonblack Female) addplot(line GGSurvivalNonblackFemale ageexitm30, sort)

Very good fit for Non-blacks
Hazard of death for non-black females

\[
\text{Graph twoway (line hazardNonblackFemale _t, sort xtitle(Years after Age 30) /// ytitle(Predicted GG hazard of death) title(Nonblack Female)) if _t<70}
\]

Fit of non-conventional GG to non-parametric survival for Black Females

\[
\text{sts graph if Female==1 & Black==1, surv xtitle(Years after Age 30) ytitle(% Alive) /// title(Black Female) addplot(line GGSurvivalBlackFemale ageexitm30, sort)}
\]
Hazard of death for black females

Relative [black:non-black] hazard of death among female

GG saturated reproduces Non-proportional Cox
Parametric Models: quantiles, not just hazards

Quantiles of GG(b,s,k) in Stata

The quantiles of any GG distribution can be easily obtained from the quantiles of a standard GG distribution with b=0, s=1, k. From section 2.1.2 of Cox/Muñoz SIM 2007, $t_{GG}(\beta,\sigma,\kappa)(p)$ the pth percentile of $GG(\beta,\sigma,\kappa)$ is

$$t_{GG}(\beta,\sigma,\kappa)(p) = e^\beta[t_{GG}(0,1,\kappa)(p)]^\sigma$$

where $t_{GG}(0,1,\kappa)(p)$, the pth percentile of $GG(0,1,\kappa)$ is

$$[\kappa^2\Gamma^{-1}(p \times (\kappa > 0) + (1-p) \times (\kappa < 0); \kappa^{-2})]^{1/\kappa}$$

*in Stata, the pth quantile of GG(b,s,k) is

$$\exp(b) \times ((k^2)^{\text{invgammap}(k^2,p*(k>0)+(1-p)*(k<0))})^{s/k}$$
Quantiles after age 30 for Black Females

```
line QuantilesBlackFemale p
```

Quantiles after age 30 for Non-black Females

```
egen samplesize= max( _n )
gen p= _n / samplesize /*trick to create a sequence of percentages deceased*/
egen QuantilesNonblackFemale=
    exp([_t]_cons)*((e(kappa)^2)*invgammap(e(kappa)^(-2),
    (e(kappa)>0)*p+(e(kappa)<=0)*(1-p)))^(e(sigma)/e(kappa))
```
**Differences# (Black - Nonblack) of Quantiles among Females**

*from section 2.1.2 of Cox/Muñoz SIM 2007*

*p*th quantile of GG(b,s,k) is

\[ \exp(b)^{((k^2)^{\text{invgammap}}(k^{-2}, p(k>0)+(1-p)(k<0)))^{s/k}} \]

*95% CI for difference in 20th

\[
\text{nlcom (twentieth: } \exp([_t\_cons+[_t]Black)* ///
((([kappa]_cons+[kappa]Black)^2)^{\text{invgammap}}(([kappa]_cons+[kappa]Black)^{-2}, 0.20)) ///
^\left(\exp([ln\_sig]_cons+[ln\_sig]Black)/([kappa]_cons+[kappa]Black) \right) ///
- \exp([_t\_cons]^2) ///
((([kappa]_cons)^2)^{\text{invgammap}}(([kappa]_cons)^{-2}, 0.20)) ///
^\left(\exp([ln\_sig]_cons)/([kappa]_cons))
\]

# except at zero, differences can't ever be constant

| _t  | Coef.  | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|-----|--------|-----------|-------|-----|---------------------|
| twentieth | -6.426571 | .8612467 | -7.46 | 0.000 | -8.114584 to -4.738558 |

---

**differences (Black - Nonblack) of Quantiles among Females**

| _t  | Coef.  | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|-----|--------|-----------|-------|-----|---------------------|
| twentieth | -6.426571 | .8612467 | -7.46 | 0.000 | -8.114584 to -4.738558 |
Differences (Black - Nonblack) of Quantiles among Females

\[ \text{gen GGDiffPercentileBtoWamongFe} = \text{QuantilesBlackFemale - QuantilesNonblackFemale} \]
\[ \text{. line GGDiffPercentileBtoWamongFe} \text{ if } p > 0.05 \text{ & } p < 0.95 \]

For the first 50%, blacks died younger (6.4 yrs for the first 20%); but for the top 25%, blacks survive longer (2 yrs for top 20%)

Summary

Analysis of racial differences in survival among women

Thumbs down: proportional hazards and conventional (proportional times) parametric models.

Thumbs up: saturated GG which reproduces Cox’s time dependent hazard ratios and more importantly, it allows for quantification of years lost/gained.

Thumbs up: Stata’s streg with anc and anc2 to incorporate non-proportionalities; and nlcom to quantify years lost/gained.
Differences (Female - Male) of Quantiles among Non-blacks

Women live longer than men, no matter at which percentile! Life expectancy (i.e., median) is 5.5 years longer.