Estudo sobre cirrose biliar primária (CBP)

Estudo realizado entre 1974 e 1984 com 418 pacientes diagnosticados com CBP. Todos foram acompanhados até a morte ou até o término do estudo. Para cada paciente tem-se:

- Tempo ti (em anos) até o óbito ou até o final do estudo
- Variável indicadora de falha δi (1 se falha e 0 se censura)
- Covariáveis: x_{i1} = idade (em anos), x_{i2} = presença de edema (0 se não e 1 se sim), x_{i3} = concentração de albumina (em g/dL), x_{i4} = concentração sérica de bilirrubina (em mg/dL) e x_{i5} = tempo de protrombina (em seg).

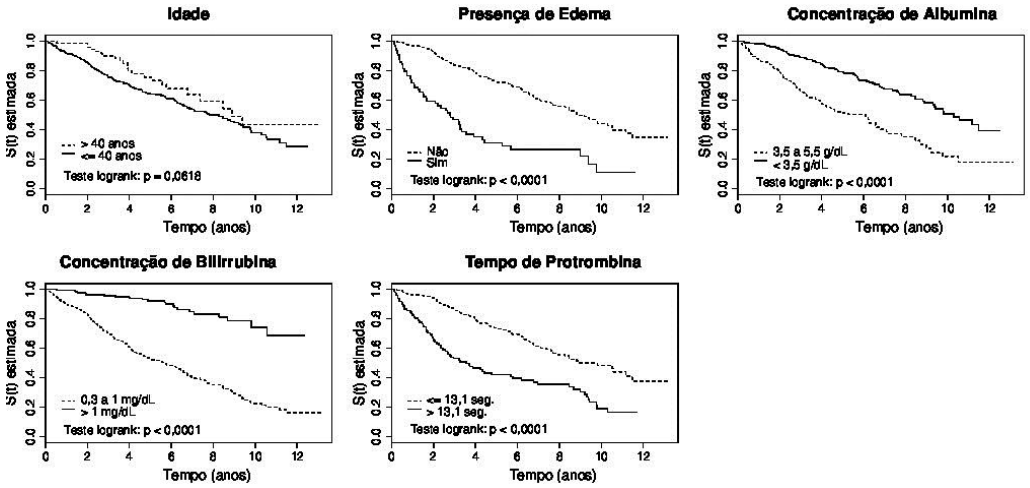
Covariáveis	Descrição
Idade	26 a 78 anos (média = 51 e d.p. = 10,47)
Edema	85% de ausência (n = 352)
Albumina	1,96 a 4,64 mg/dL (média = 3,5 e d.p. = 0,42)
Bilirrubina	0,3 a 28 mg/dL (média = 3,2 e d.p. = 4,38)
Protrombina	9 a 18 segundos (média = 10,73 e d.p. = 1,02)

Para a análise dos dados, as três últimas covariáveis foram consideradas na escala logarítmica. Além disso, as covariáveis contínuas foram centradas em seus respectivos valores médios e dois pacientes foram excluídos devido à presença de dados faltantes nas covariáveis. Ao final do estudo foi registrado 55% de observações censuradas.

PREPARANDO OS DADOS

```
library(survival)
library(timereg)
library(survivalROC)

data(pbc); names(pbc); attach(pbc); str(pbc)
i<-order(pbc[,2])
pbc1<-pbc[i,]
dim(pbc1); detach(pbc); attach(pbc1)
dat1<-as.data.frame(cbind(age, albumin, bili, edema, protime, time, status))
pbc2<-na.omit(dat1)
dim(pbc2); detach(pbc1); attach(pbc2)

idade<-age- mean(age)                          # idade centrada na média
logalbum<-log(albumin)-mean(log(albumin))  # logalbum centrada na média
logbili<-log(bili)-mean(log(bili))          # logbili centrada na média
logprotime<-log(protime)-mean(log(protime))# logprotime centrada na média
Status<-ifelse(status==0, 0, 1)             # indicadora de censura
Edema<-ifelse(edema==0, 0, 1)               # edema em 2 categorias
tempo<-time/365                             # tempo em anos

pbc3<-as.data.frame(cbind(idade,logalbum,logbili,Edema,logprotime,tempo,Status))
detach(pbc2); attach(pbc3)

set.seed(157)
n1<-dim(pbc3)[1]       # n1 = número de linhas do banco de dados
ei<-rnorm(n1,0,0.001)
tempos<- tempo + ei    # somando delta pequeno para evitar tempos empatados
pbc3$tempos<-tempos
```

ANÁLISE EXPLORATÓRIA



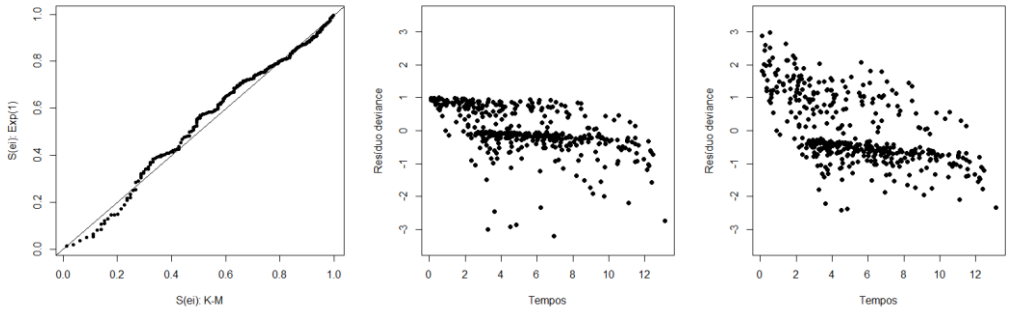
Ajuste do Modelo de Cox padrão

```
mod1<-coxph(Surv(tempos,Status)~idade+logalbum+logbili+Edema+logprotime, method="breslow",
                                                                        data=pbc3)
```

```
               coef exp(coef) se(coef)     z       p
idade        0.0232    1.0234  0.00722  3.21 1.3e-03
logalbum    -2.6808    0.0685  0.60202 -4.45 8.5e-06
logbili      0.8461    2.3305  0.07637 11.08 0.0e+00
Edema        0.4836    1.6219  0.18355  2.63 8.4e-03
logprotime   1.7764    5.9086  0.71632  2.48 1.3e-02
```

```
# Avaliando a Suposição de Riscos Proporcionais
cox.zph(mod1, transform="log")
              rho  chisq        p
idade      -0.0675  0.789 0.374438
logalbum    0.0769  1.181 0.277239
logbili     0.0990  1.515 0.218397
Edema      -0.1543  4.295 0.038226
logprotime -0.3239 12.692 0.000367
GLOBAL          NA 25.428 0.000115
```

Verificando se os resíduos de Cox-Snell ~ Exp(1) e investigando os resíduos martingal e deviance



Qualidade de predição do modelo de Cox padrão via AUC(t)

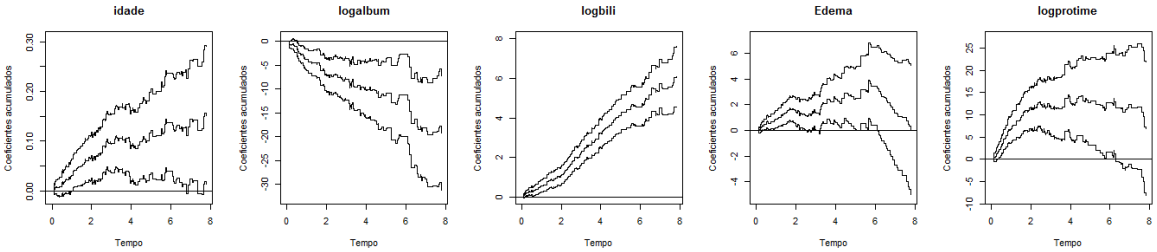
```
          NNE        KM   t    t max
[1,] 0.9238957 0.9364686 0.5 13.13622
[2,] 0.8514080 0.8681287 1.5 13.13622
[3,] 0.8484198 0.8751338 3.0 13.13622
[4,] 0.8415698 0.8840175 5.0 13.13622
[5,] 0.8188024 0.8567686 7.0 13.13622
[6,] 0.7919631 0.8312938 8.0 13.13622
```

Ajuste do Modelo de Cox Dinâmico

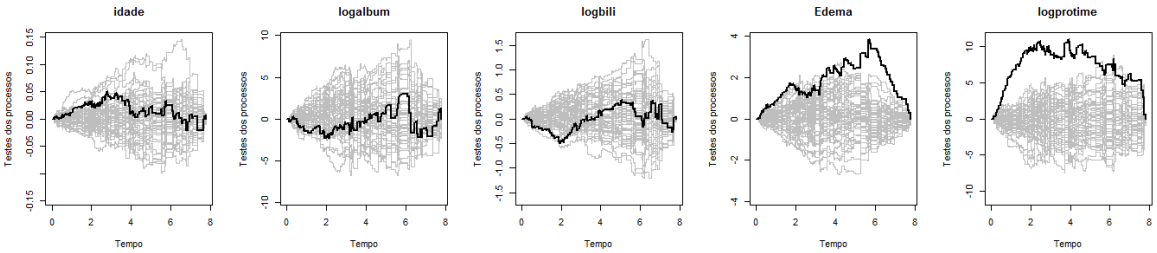
```
mod2<-timecox(Surv(tempos, Status) ~ idade + logalbum + logbili + Edema +
              logprotime, max.time=8, residuals=1, data=pbc3)
summary(mod2)
```

```
Test for non-significant effects
             Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                          39.60                0.000
idade                                 3.83                0.001
logalbum                              5.61                0.000
logbili                              10.10                0.000
Edema                                 3.78                0.002
logprotime                            6.28                0.000

Test for time invariant effects
                  Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                        1.9900                       0.000
idade                              0.0503                       0.612
logalbum                           2.9400                       0.921
logbili                            0.5060                       0.740
Edema                              3.8500                       0.000
logprotime                        11.0000                       0.002
                    Cramer von Mises test p-value H_0:constant effect
(Intercept)                     1.66e+01                       0.000
idade                           3.83e-03                       0.592
logalbum                        1.18e+01                       0.896
logbili                         3.98e-01                       0.689
Edema                           3.14e+01                       0.000
logprotime                      4.55e+02                       0.000
```

```
par(mfrow=c(1,5))
plot(mod2, ylab="Coeficientes acumulados", xlab="Tempo", specific.comps=2)
plot(mod2, ylab="Coeficientes acumulados", xlab="Tempo", specific.comps=3)
plot(mod2, ylab="Coeficientes acumulados", xlab="Tempo", specific.comps=4)
plot(mod2, ylab="Coeficientes acumulados", xlab="Tempo", specific.comps=5)
plot(mod2, ylab="Coeficientes acumulados", xlab="Tempo", specific.comps=6)
```




```
par(mfrow=c(1,5))
plot(mod2, score=T, xlab="Tempo", ylab="Testes dos processos",specific.comps=2)
plot(mod2, score=T, xlab="Tempo", ylab="Testes dos processos",specific.comps=3)
plot(mod2, score=T, xlab="Tempo", ylab="Testes dos processos",specific.comps=4)
plot(mod2, score=T, xlab="Tempo", ylab="Testes dos processos",specific.comps=5)
plot(mod2, score=T, xlab="Tempo", ylab="Testes dos processos",specific.comps=6)
```

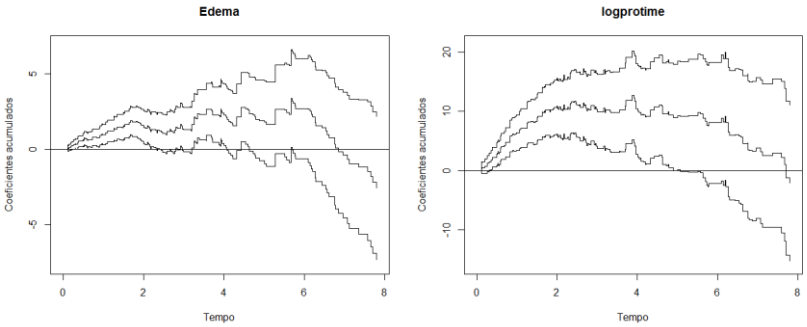


Ajuste do Modelo de Riscos Multiplicativos Semiparamétrico

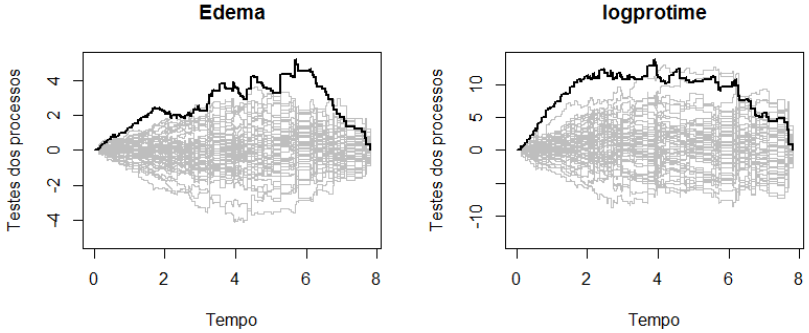
```
mod2.1<-timecox(Surv(tempos,Status) ~ const(idade) + const(logalbum) +
        const(logbili) + Edema + logprotime, residuals=1, max.time=8,
        data=pbc3)
summary(mod2.1)
```

```
Multiplicative Hazard Model
Test for nonparametric terms
Test for non-significant effects
             Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                          26.70                0.000
Edema                                 4.03                0.003
logprotime                            5.95                0.000
Test for time invariant effects
                  Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                          4.18                         0.003
Edema                                5.21                         0.001
logprotime                          13.80                         0.002
                    Cramer von Mises test p-value H_0:constant effect
(Intercept)                          72.2                         0.000
Edema                                66.0                         0.002
logprotime                          694.0                         0.000
Parametric terms:
                 Coef.    SE Robust SE     z P-val
const(idade)     0.019 0.008     0.010  1.90 0.058
const(logalbum) -2.750 0.611     0.664 -4.15 0.000
const(logbili)   0.767 0.088     0.084  9.13 0.000
```

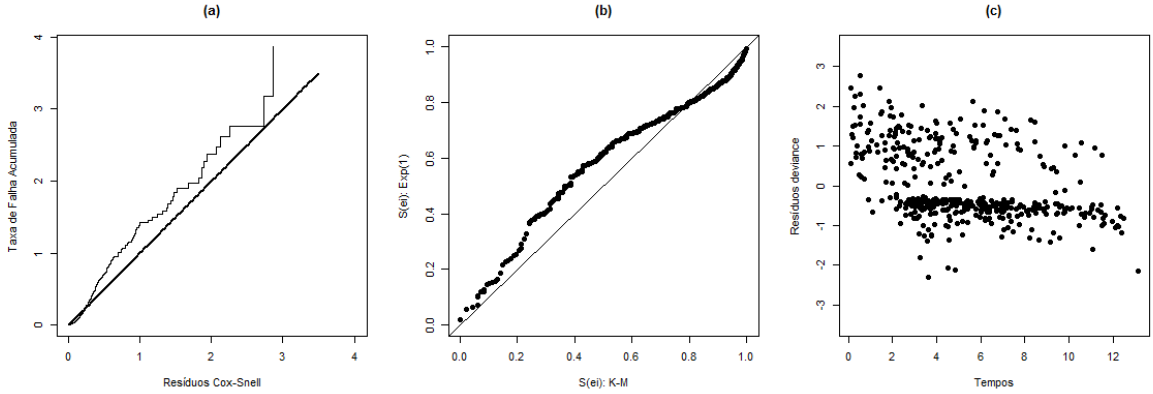
```
par(mfrow=c(1,2))
plot(mod2.1, ylab="Coeficientes acumulados", xlab="Tempo", specific.comps=2)
plot(mod2.1, ylab="Coeficientes acumulados", xlab="Tempo", specific.comps=3)
```




```
par(mfrow=c(1,2))
plot(mod2.1, score=T, xlab="Tempo", ylab="Testes dos processos",specific.comps=2)
plot(mod2.1, score=T, xlab="Tempo", ylab="Testes dos processos",specific.comps=3)
```



QUALIDADE DE AJUSTE DO MODELO – ANÁLISE DE RESÍDUOS




```
mod2.1<-timecox(Surv(tempos, Status) ~ const(idade) + const(logalbum) + const(logbili)+
                Edema + logprotime, residuals=1, data=pbc3)

n<-dim(pbc3)[1]
rm<-matrix(0,n,1)                                   # resíduos martingal
for(i in 1:n){
 rm[i]<-sum(mod2.1$residuals$dM[,i])
}

delta<-pbc3$Status
ei<-delta-rm                                        # resíduos Cox-Snell

rd <- sign(rm)*sqrt(-2*(rm + delta*log(abs(ei))))   # resíduos deviance

## Gráficos qualidade global

r.surv <- survfit(Surv(ei, delta)~1)
e<-r.surv$time
He<- -log(r.surv$surv)

par(mfrow=c(1,3))
plot(e, He, type="s",xlab="Resíduos Cox-Snell",ylab="Taxa de Falha Acumulada")
t <- seq(0, 3.5,length=100)
lines(t,t,lwd=2)
title("(a)", cex=1)

st<- r.surv$surv
sexp<-exp(-e)
plot(st,sexp,xlab="S(ei): K-M",ylab="S(ei): Exp(1)", pch=16,ylim=c(0,1),xlim=c(0,1))
abline(a=0,b=1,lwd=1)
title("(b)", cex=1)

ti<-pbc3$tempos
plot(ti,rd, pch=16, xlab="Tempos", ylab="Resíduos deviance", ylim=c(-3.5,3.5))
title("(c)", cex=1)
```

AUC

	NNE	KM	t	t_max
[1,]	0.9189124	0.9544142	0.5	13.13622
[2,]	0.8831921	0.8838450	1.5	13.13622
[3,]	0.8708101	0.8736829	3.0	13.13622
[4,]	0.8754957	0.8836529	5.0	13.13622
[5,]	0.8574723	0.8596613	7.0	13.13622
[6,]	0.8187665	0.8385842	8.0	13.13622

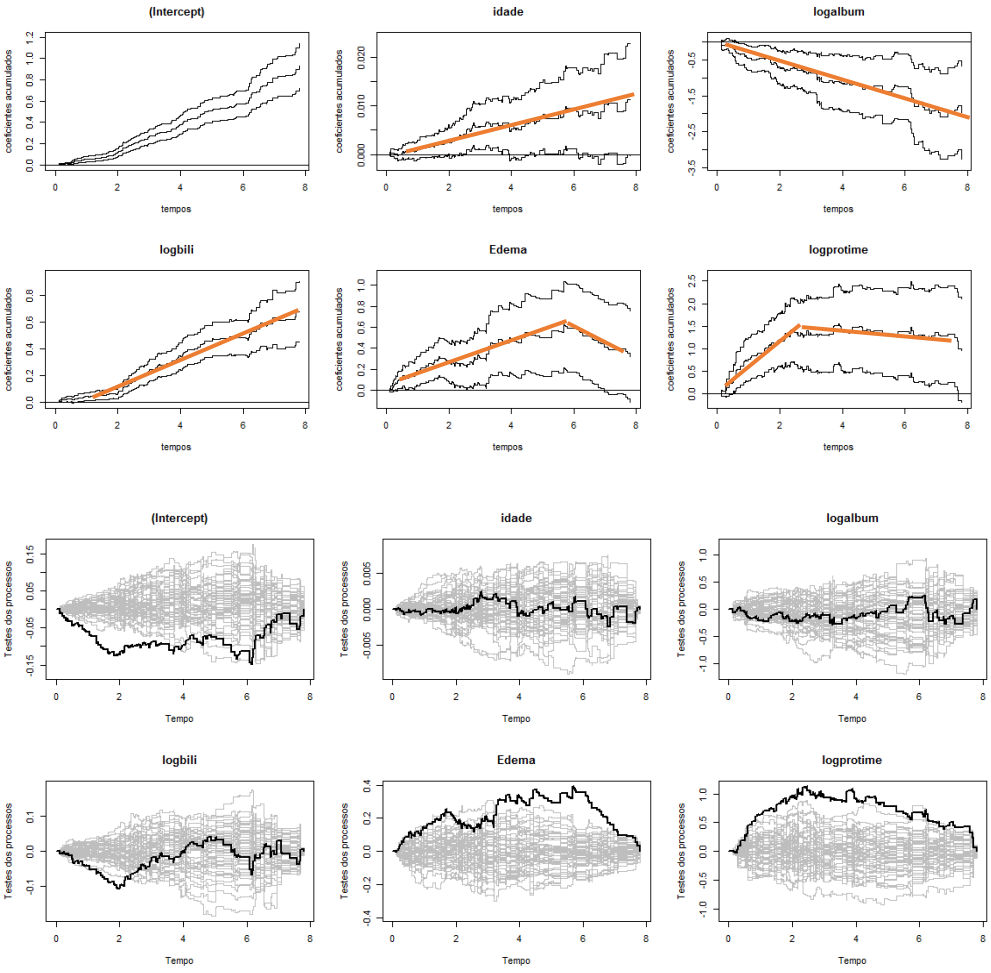
Ajuste do Modelo Aditivo de Aalen

```
mod3<- aalen(Surv(tempos, Status) ~ idade + logalbum + logbili + Edema +
             logprotime, residuals=1, max.time=8, data=pbc3)
summary(mod3)
```

	Supremum-test of significance	p-value H_0: B(t)=0
(Intercept)	10.50	0.000
idade	2.78	0.086
logalbum	3.88	0.006
logbili	7.67	0.000
Edema	3.50	0.012
logprotime	3.49	0.005

Test for time invariant effects

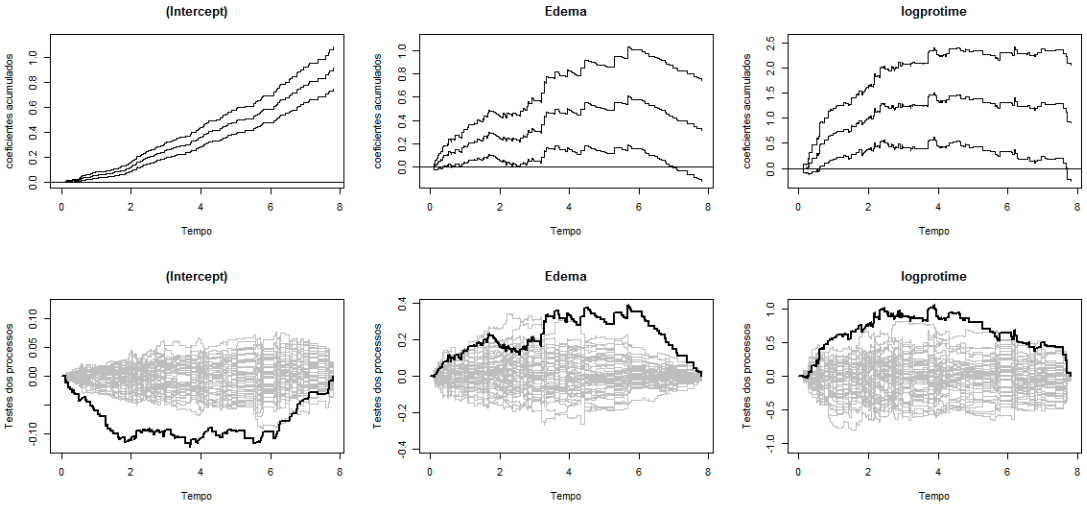
	Kolmogorov-Smirnov test	p-value H_0:constant effect
(Intercept)	0.14800	0.059
idade	0.00244	0.967
logalbum	0.28100	0.955
logbili	0.10800	0.283
Edema	0.39100	0.001
logprotime	1.13000	0.001
	Cramer von Mises test	**p-value H_0:constant effect**
(Intercept)	6.40e-02	0.022
idade	6.20e-06	0.982
logalbum	1.52e-01	0.868
logbili	1.46e-02	0.379
Edema	4.39e-01	0.000
logprotime	4.24e+00	0.000



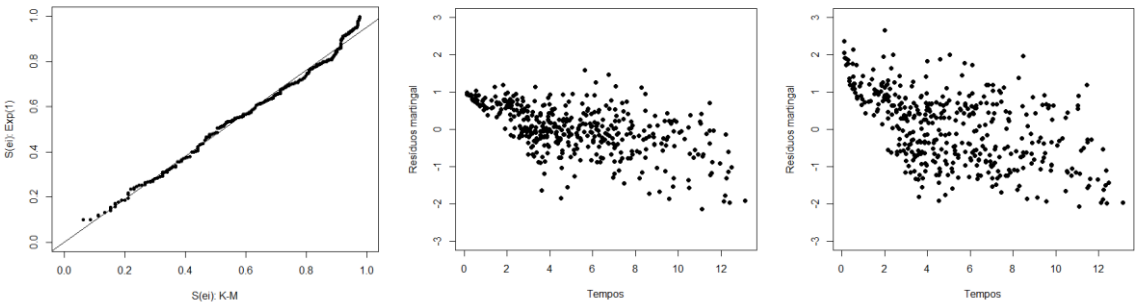
Ajuste do Modelo de Riscos Aditivos Semiparamétrico

```
mod3.1<- aalen(Surv(tempos, Status) ~ const(idade) + const(logalbum) +
  const(logbili) + Edema + logprotime, residuals=1, max.time=8, data=pbc3)
summary(mod3.1)
```

```
Supremum-test of significance p-value                H_0: B(t)=0
(Intercept)                           12.30                 0.000
Edema                                  3.31                 0.027
logprotime                             3.18                 0.026
Test for time invariant effects
                   Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                          0.123                       0.002
Edema                                0.387                       0.002
logprotime                           1.060                       0.003
                     Cramer von Mises test p-value H_0:constant effect
(Intercept)                         0.0623                       0.000
Edema                               0.4390                       0.000
logprotime                          3.7100                       0.001
Parametric terms:
                 Coef.    SE Robust SE     z P-val
const(idade)     0.001 0.001     0.001  2.27 0.023
const(logalbum) -0.269 0.077     0.068 -3.94 0.000
const(logbili)   0.081 0.010     0.009  9.12 0.000
```



Verificando se os resíduos de Cox-Snell ~ Exp(1) / Resíduos martingale e deviance

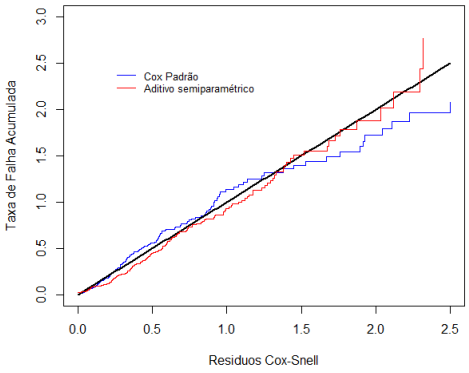


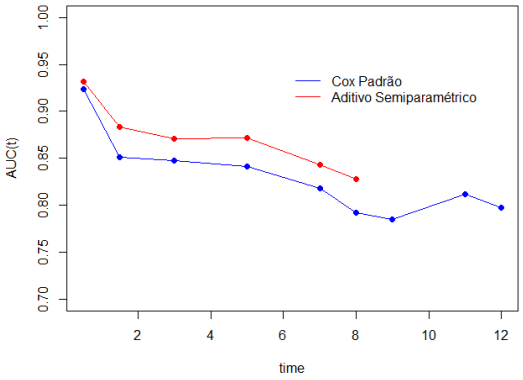
Qualidade de predição do modelo de Riscos Aditivos Semiparamétrico - AUC(t)

AUC

```
          NNE        KM   t    t max
[1,] 0.9320453 0.9527640 0.5 13.13622
[2,] 0.8826027 0.8839181 1.5 13.13622
[3,] 0.8709690 0.8719288 3.0 13.13622
[4,] 0.8725529 0.8797192 5.0 13.13622
[5,] 0.8426008 0.8557122 7.0 13.13622
[6,] 0.8285650 0.8381898 8.0 13.13622
```

Comparando o ajuste global dos modelos de Cox Padrão e de riscos aditivo semiparamétrico





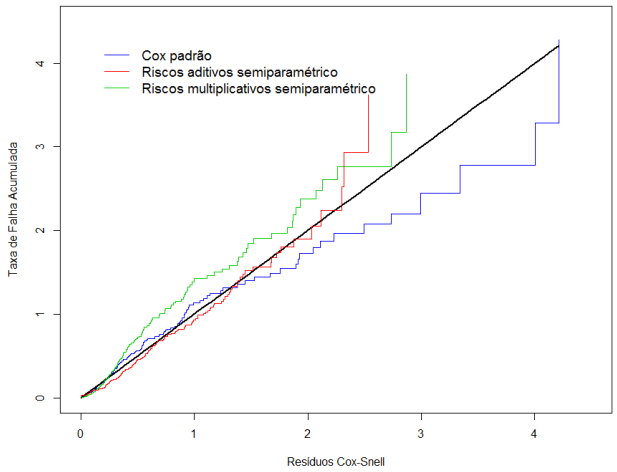
Cox

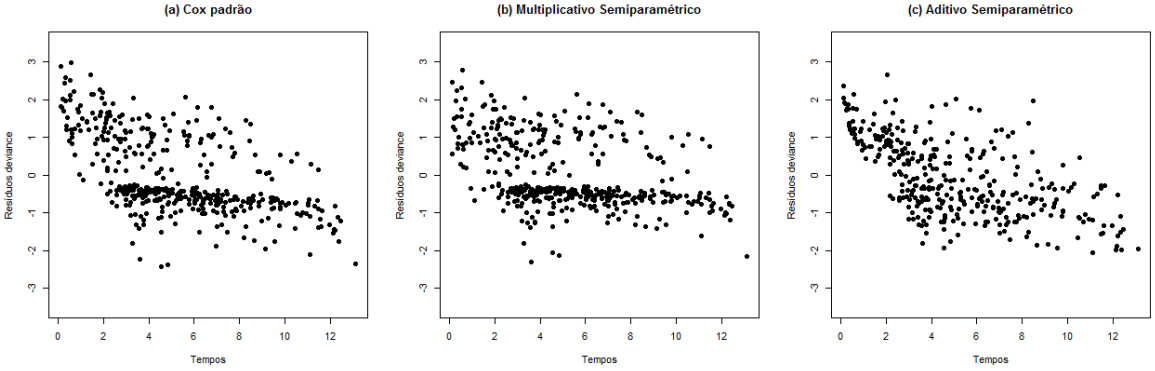
```
      NNE    t    t max
[1,] 0.924  0.5 13.13622
[2,] 0.851  1.5 13.13622
[3,] 0.848  3.0 13.13622
[4,] 0.841  5.0 13.13622
[5,] 0.818  7.0 13.13622
[6,] 0.792  8.0 13.13622
[7,] 0,785  9,0 13.13622
[8,] 0,812 11,0 13.13622
[9,] 0,798 12,0 13.13622
```

Riscos Aditivos Semiparamétrico

```
      NNE    t    t max
[1,] 0.932  0.5 13.13622
[2,] 0.883  1.5 13.13622
[3,] 0.871  3.0 13.13622
[4,] 0.872  5.0 13.13622
[5,] 0.843  7.0 13.13622
[6,] 0.828  8.0 13.13622
 --
 --
 --
```

Comparação dos três Modelos

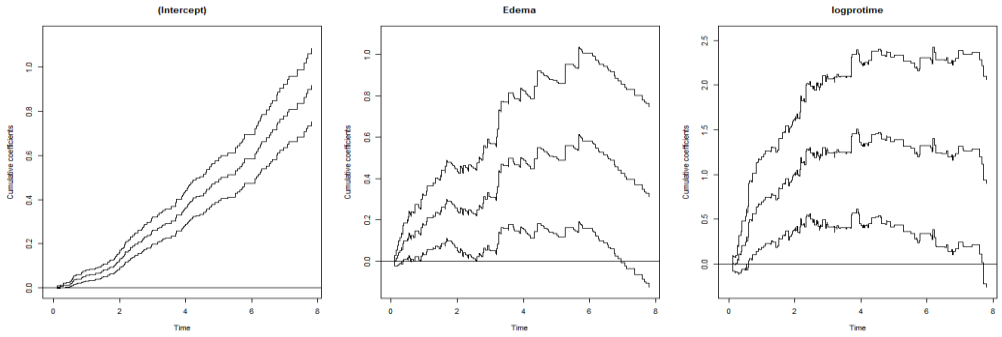




Modelo Selecionado – Riscos Aditivos Semiparamétrico

```
mod3.1<- aalen(Surv(tempos, Status) ~ const(idade) + const(logalbum) +
          const(logbili) + Edema + logprotime, residuals=1, max.time=8,
          data=pbc3)
summary(mod3.1)
```

```
Supremum-test of significance p-value                H_0: B(t)=0
(Intercept)                            12.30               0.000
Edema                                   3.31               0.027
logprotime                              3.18               0.026
Test for time invariant effects
                  Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                            0.123                   0.002
Edema                                  0.387                   0.002
logprotime                             1.060                   0.003
                    Cramer von Mises test p-value H_0:constant effect
(Intercept)                           0.0623                   0.000
Edema                                 0.4390                   0.000
logprotime                            3.7100                   0.001
Parametric terms:
                 Coef.    SE Robust SE     z P-val
const(idade)     0.001 0.001     0.001  2.27 0.023
const(logalbum) -0.269 0.077     0.068 -3.94 0.000
const(logbili)   0.081 0.010     0.009  9.12 0.000
```



$$\widehat{\Lambda}(t\,|x_i) = \widehat{B}_0(t) + (\widehat{\beta}_1 * t)(x_{i1} - \bar{x}_1) + \widehat{B}_2(t)\,x_{i2} + (\widehat{\beta}_3 * t)(x_{i3} - \bar{x}_3) + (\widehat{\beta}_4 * t)(x_{i4} - \bar{x}_4) + \widehat{B}_5(t)(x_{i5} - \bar{x}_5)$$

em que x_{i1} = idade (em anos) e $\bar{x}_1$ = média das idades, x_{i2} = presença de edema (1 = sim e 0 = não),
x_{i3} = ln(albumina) e $\bar{x}_3$ = média de ln(albumina), x_{i4} = ln(bilirrubina) e $\bar{x}_4$ = média de ln(bilirrubina),
x_{i5} = ln(tempo de protrombina) e $\bar{x}_5$ = média de ln(tempo de protrombina).

```
> mod3.1$cum
            time (Intercept)       Edema   logprotime
[115,] 4.0043559 0.364485533 0.488951379  1.331108392
[133,] 5.0068285 0.493053386 0.494809893  1.385191238
[160,] 7.0766349 0.797373148 0.409019964  1.274976578
```

Edema 1 | 0 mantida as demais constantes no valor médio

t = 4 → $RR(t=4) = \frac{\widehat{\Lambda}(t\,|x_i)}{\widehat{\Lambda}(t\,|x_j)} = \frac{\hat{B}_0(t)+(\widehat{\beta}_1*t)*0+\hat{B}_2(t)*1+(\widehat{\beta}_3*t)*0+(\widehat{\beta}_4*t)*0+\hat{B}_5(t)*0}{\hat{B}_0(t)+(\widehat{\beta}_1*t)*0+\hat{B}_2(t)*0+(\widehat{\beta}_3*t)*0+(\widehat{\beta}_4*t)*0+\hat{B}_5(t)*0}$

$= \frac{0.3645+0.4889}{0.3645} = 2.341289$

t = 5 → $RR(t=5) = \frac{0.4930+0.4948}{0.4930} = 2.003651$

t = 7 → $RR(t=7) = \frac{0.7974+0.4090}{0.7974} = 1.51291$

idade 66 | 46 mantida as demais constantes no valor médio e edema = 0

```
> mean(pbc2$age)
[1] 50.76
```

Parametric terms:

	Coef.	SE	Robust SE	z	P-val
const(idade)	0.001	0.001	0.001	2.27	0.023
const(logalbum)	-0.269	0.077	0.068	-3.94	0.000
const(logbili)	0.081	0.010	0.009	9.12	0.000

```
> mod3.1$cum
            time (Intercept)       Edema   logprotime
[115,] 4.0043559 0.364485533 0.488951379  1.331108392
[133,] 5.0068285 0.493053386 0.494809893  1.385191238
[160,] 7.0766349 0.797373148 0.409019964  1.274976578
```

$t = 4 \rightarrow RR(t=4) = \frac{\hat{\lambda}(t|x_i)}{\hat{\lambda}(t|x_j)} = \frac{\hat{B}_0(t) + (\hat{\beta}_1 * t) * (66 - 50.76)}{\hat{B}_0(t) + (\hat{\beta}_1 * t) * (56 - 50.76)} = \frac{\hat{B}_0(t) + (\hat{\beta}_1 * t) * (15{,}24)}{\hat{B}_0(t) + (\hat{\beta}_1 * t) * (-4{,}76)}$

$= \frac{0.3645 + (0.001 * 4) * (15{,}24)}{0.3645 + (0.001 * 4) * (-4.76)} = 1.148$

$t = 5 \rightarrow RR(t=5) = \frac{0.4930 + (0.001 * 5) * (15{,}24)}{0.4930 + (0.001 * 5) * (-4.76)} = 1.185$

$t = 7 \rightarrow RR(t=7) = \frac{0.7974 + (0.001 * 7) * (15{,}24)}{0.7974 + (0.001 * 7) * (-4.76)} = 1.259$

Logo, o risco de óbito do indivíduo i é ≈ 1,2 vezes o do indivíduo j para qualquer t, tendo em vista que o efeito da idade é constante ao longo do tempo.