MAYO CLINIC

MULTISTATE ANALYSIS USING {SURVIVAL}

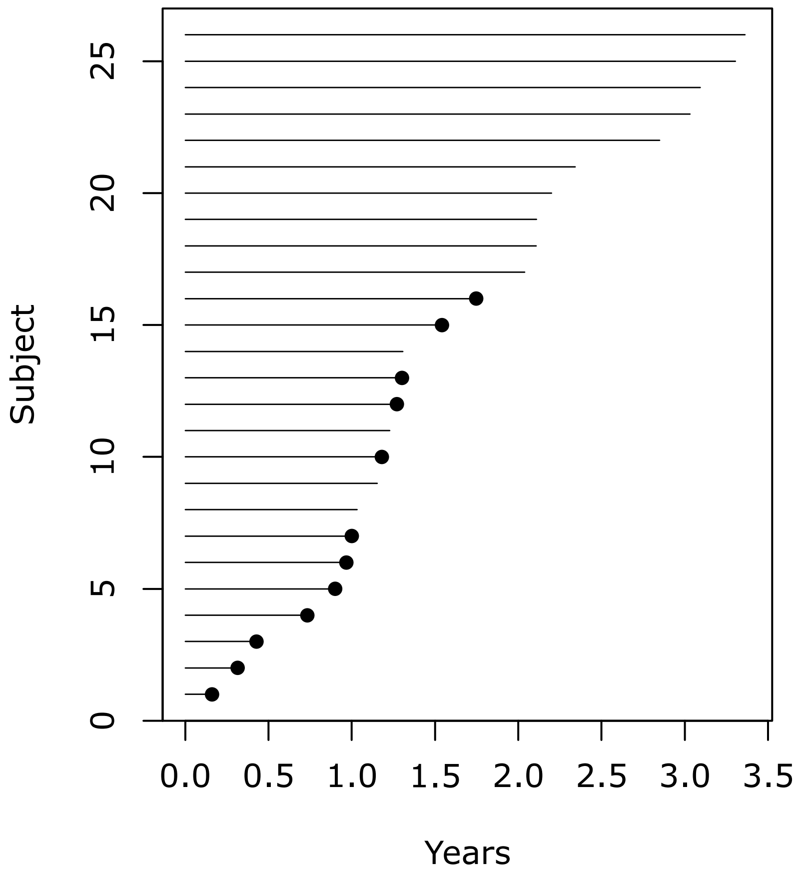
Elizabeth J. Atkinson

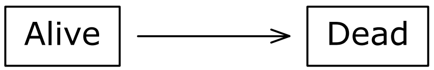
R/Medicine
8/26/2021

CLASSICAL SURVIVAL ANALYSIS

TYPE OF DATA

- boxes (states) and arrows (transitions)



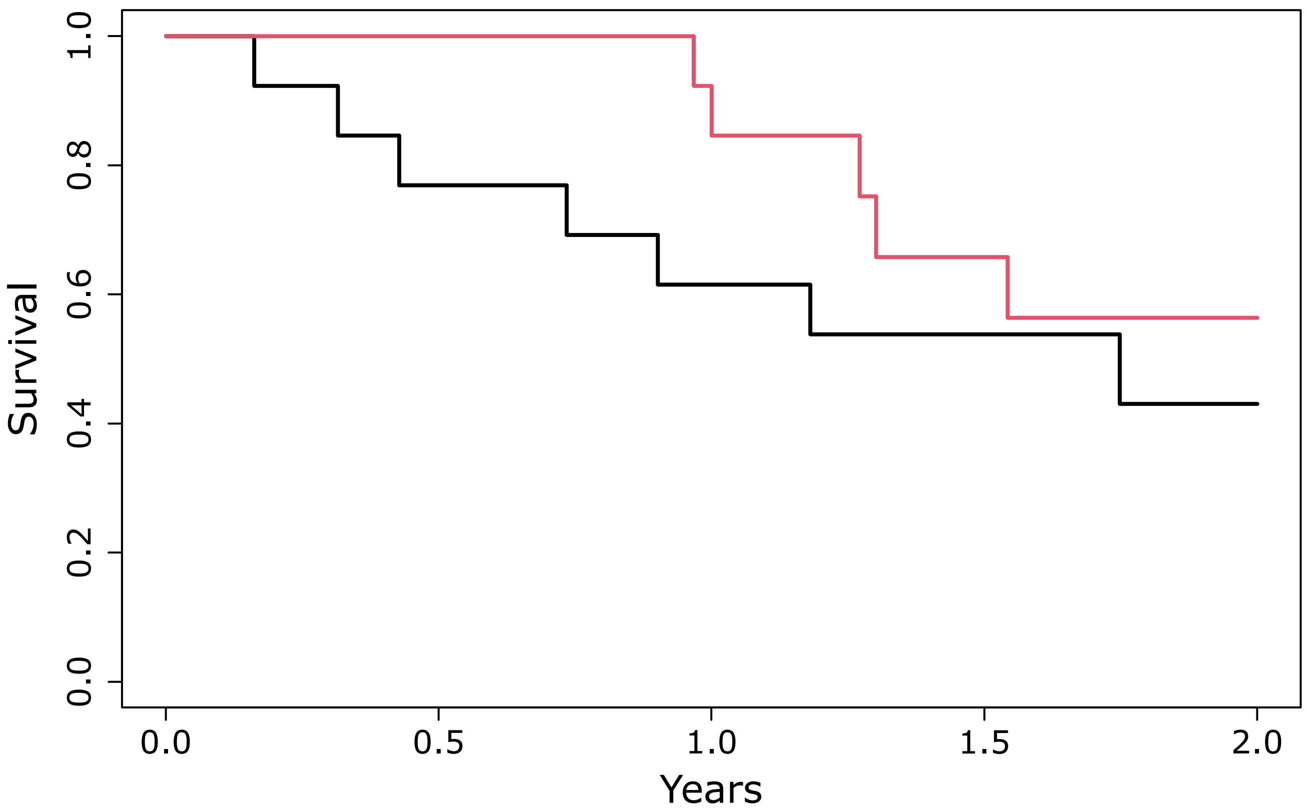


DATA

- 1 row per subject: `Surv(time, status)`
- Time-dependent covariates, multiple events of the same type
  - counting process format: `Surv(time1, time2, status)`
- status values: 0/1, FALSE/TRUE, 1/2

MAIN TOOLS

- Kaplan-Meier curves: survfit()



MAIN TOOLS

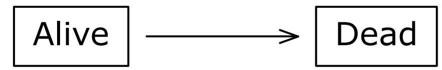
- Cox model: coxph()

```
Call:
coxph(formula = Surv(futime, fustat) ~ rx,
      data = ovarian)

      coef exp(coef) se(coef)      z    p
rx -0.5964    0.5508   0.5870 -1.016 0.31

Likelihood ratio test=1.05  on 1 df, p=0.3052
n= 26, number of events= 12
```

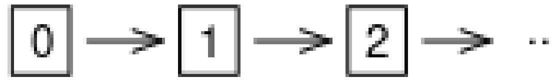

QUESTIONS

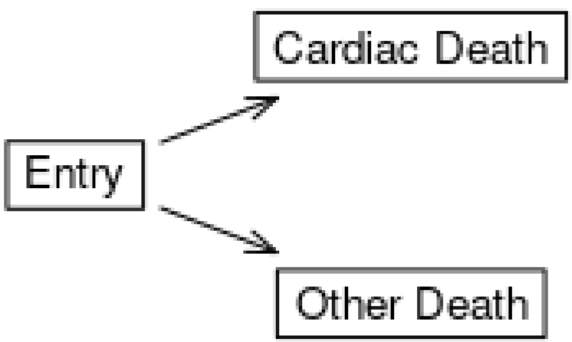


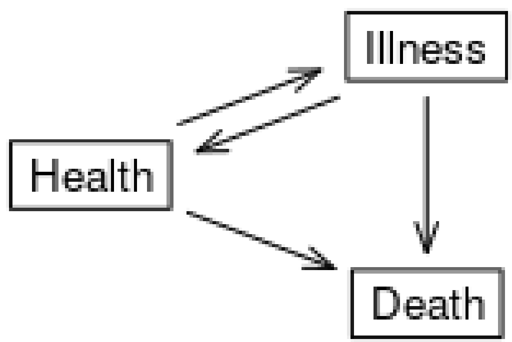
- What is the probability of being in the 1st state at time *t*?
  - Kaplan-Meier, 5-year survival
- How long does it take to reach the 2nd state?
  - median survival
- How long does someone remain in the 1st state?
  - restricted mean time in state (RMTS)
- What is the risk of moving to the 2nd state? (arrow)
  - hazard model (Cox)

MULTISTATE ANALYSIS

MULTISTATE MODELS







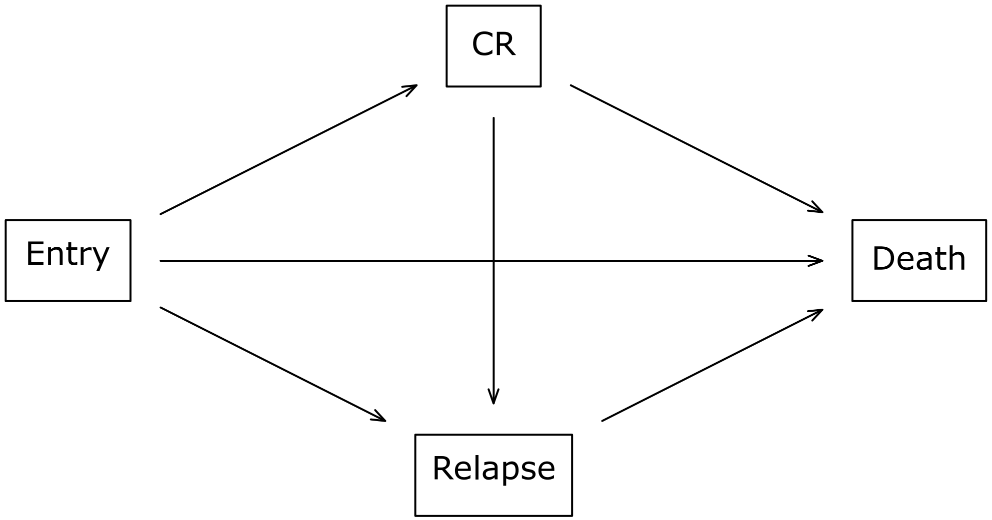
EXAMPLE

- Acute myeloid leukemia (myeloid)
- Trial with two treatments (A and B), 646 subjects
- Outcomes: complete response (CR), relapse, and death.

id	trt	sex	futime	death	crtime	rltime
1	B	f	235	1	44	113
2	A	m	286	1	NA	NA
3	A	f	1983	0	38	NA
4	B	f	2137	0	25	NA
5	B	f	326	1	56	200
6	B	f	2041	0	NA	NA

DIAGRAM

```
> states <- c("Entry", "CR", "Relapse", "Death")
> cmat <- matrix(0L,4,4, dimnames=list(states, states))
> cmat[1,2:4]  <- 1; cmat[2,3:4] <- 1; cmat[3,4] <- 1
> statefig(c(1,2,1), cmat)
```



 |

DATA REQUIREMENTS

- Subject id
- Time interval (time1, time2]
- Covariates that apply over the time interval
- Transition state, if any, at the end of the interval.
  - *factor*: 1st level is “no transition”
- Optional: Current state at time1

 |

BUILD DATA

```
> m1 <- tmerge(myeloid[,1:3], myeloid, id=id,
               dead=event(futime, death),
               cr=event(crtime),
               relapse=event(rltime))
> m1$event <- with(m1,factor(1*cr + 2*relapse +3*dead,
                     0:3, c("censor", "CR",
                            "relapse", "death")))

> m1[1:4,]
```

id	trt	sex	tstart	tstop	cr	relapse	dead	event
1	B	f	0	44	1	0	0	CR
1	B	f	44	113	0	1	0	relapse
1	B	f	113	235	0	0	1	death
2	A	m	0	286	0	0	1	death

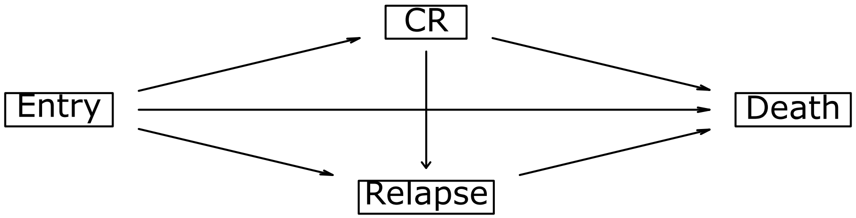
BUILD DATA

```
> summary(m1)

        early late gap within boundary leading trailing
death       0    0   0      0        0       0      646
cr          0    0   0    454        0       0        0
relapse     0    0   0    226        0       0        0
        tied missid
death      0      0
cr         0      0
relapse    0      0
```

 |

CHECK DATA



- Make sure to compare results with diagram

```
> ck <- survcheck(Surv(tstart, tstop, event) ~1,
           data=m1, id=id)

> ck$transitions
          to
from        CR relapse death (censored)
  (s0)     454      20   102         70
  CR         0     206    50        198
  relapse    0       0   168         58
  death      0       0     0          0
```

CHECK DATA

Subjects:

- cannot in two places at once (no overlapping intervals)
- cannot spend time nowhere (no time gaps)
- cannot end in one state and begin in another (teleport)

```
ck$flag
  overlap   gap   jump   teleport duplicate
        0     0      0          0         0
```


CHECK DATA

Number of repeat events:

- 151 subjects had 3 different events
- No subject experienced the same event more than once

```
ck$events
             count
state           0   1   2   3
  CR          192 454   0   0
  relapse     420 226   0   0
  death       326 320   0   0
  (any)        70 303 122 151
```


QUESTIONS

- What is the probability of being in a state at time *t*?
- How long does it take to reach a state?
- How long does someone remain in a state?
- What is the risk of moving to a state? (arrows)

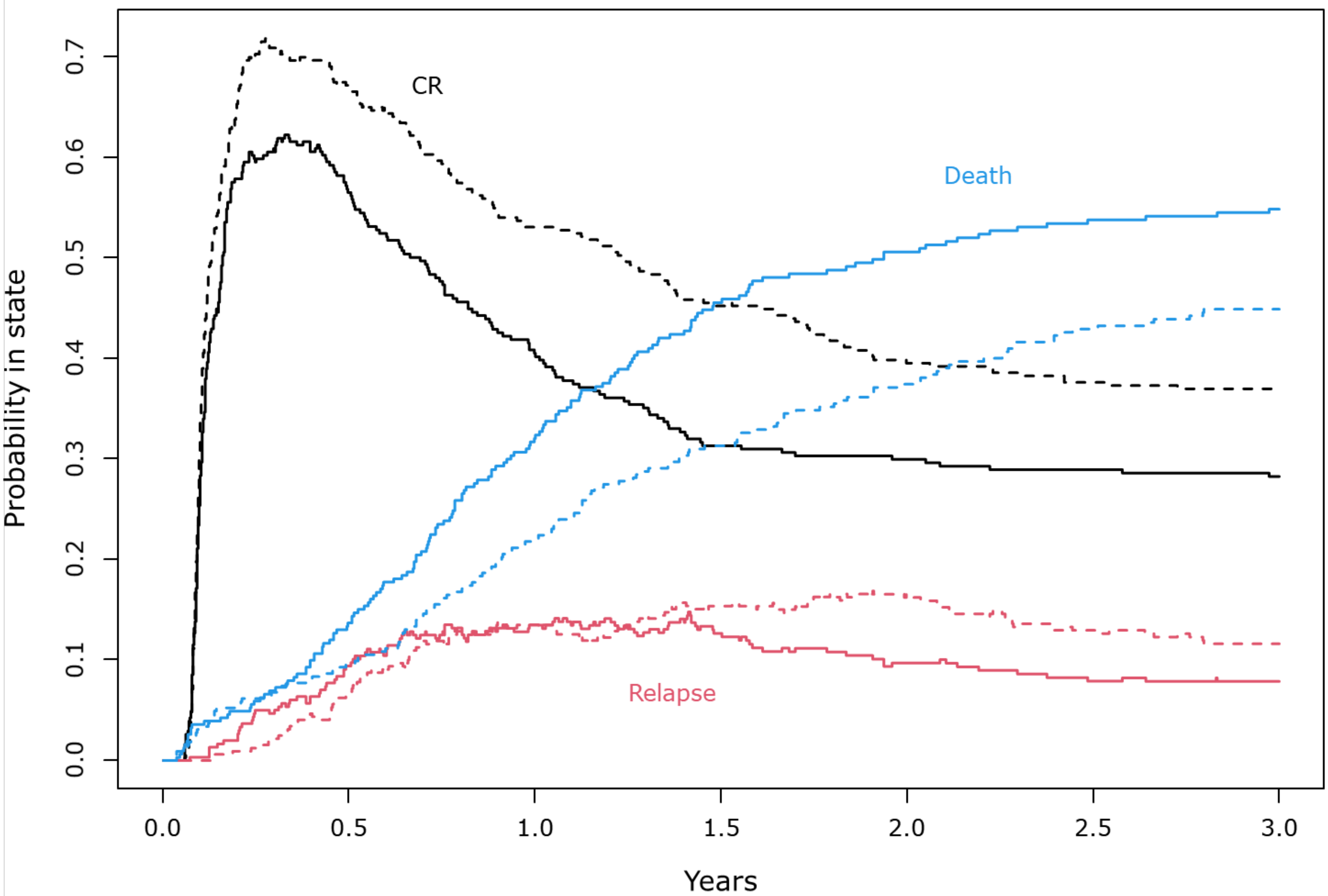
PROBABILITY-IN-STATE

```
> fit1 <- survfit(Surv(tstart, tstop, event) ~ trt,
                   data=m1, id=id)

> plot(fit1, lty=1:2, col= c(1,1, 2,2, 4,4),
       xmax= 3*365, xscale=365,
       xlab="Years", ylab="Probability in state")

> text(c(260, 500, 800), c(.67, .06, .58),
       c("CR", "Relapse", "Death"),
       col=c(1,2,4))
```

 |

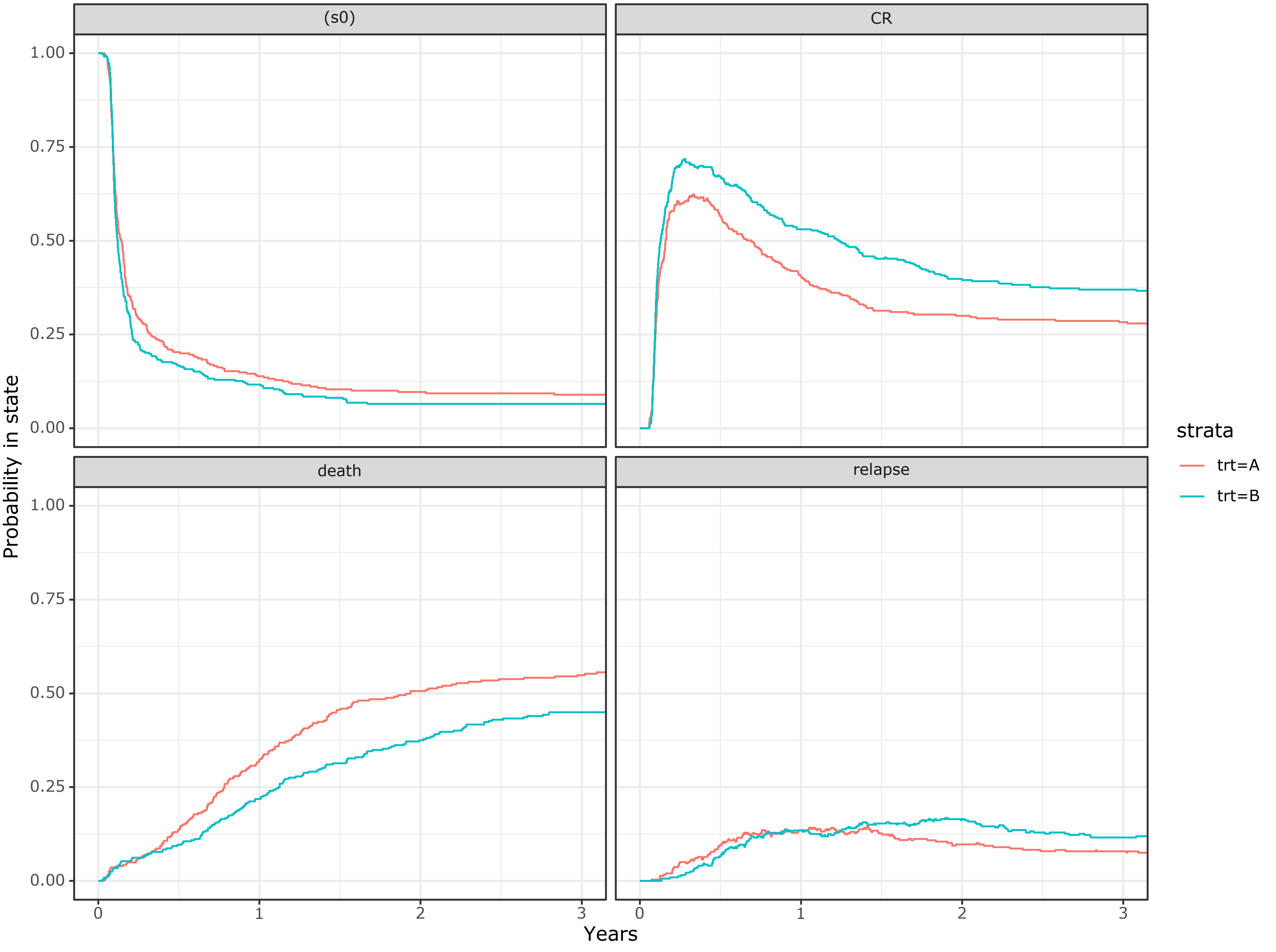


CR
Death
Relapse
Probability in state
Years
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.0
0.5
1.0
1.5
2.0
2.5
3.0

GGPLOT

```
> sdat1 <- survfit0(fit1) %>% broom::tidy()

> g1 <- ggplot(sdat1, aes(x=time/365, y=estimate,
                          color=strata)) +
  geom_step() + facet_wrap(~state) +
  coord_cartesian(xlim=c(0,3), ylim=c(0,1)) +
  xlab("Years") + ylab("Probability in state") +
  theme_bw()
```


(s0)
CR
death
relapse
1.00
0.75
0.50
0.25
0.00
0
1
2
3
Probability in state
Years
strata
trt=A
trt=B

TIME IN STATE

```
print(fit1, rmean=3*365, scale=365, digits=2)
```

```
                    n nevent rmean*
trt=A, (s0)       632      0   0.50
trt=B, (s0)       694      0   0.41
trt=A, CR         632    206   1.09
trt=B, CR         694    248   1.41
trt=A, relapse    632    109   0.29
trt=B, relapse    694    117   0.35
trt=A, death      632    171   1.12
trt=B, death      694    149   0.84
   *restricted mean time in state
    (max time = 3 )
```


FIT MULTISTATE MODELS

```
> cfit <- coxph(Surv(tstart, tstop, event) ~ trt +
            sex, data= m1, id=id)

> round(coef(cfit, matrix=TRUE), 3)

        1:2    1:3    2:3    1:4    2:4    3:4
trtB 0.215 -0.521 -0.162 -0.097 -0.653 -0.300
sexm 0.077  1.156 -0.234  0.352  0.092  0.201

attr(,"states")
[1] "(s0)"    "CR"      "relapse" "death"
```

FIT MULTISTATE MODELS

```
> cfit


1:2         coef exp(coef) se(coef) robust se     z      p
   trtB 0.21531   1.24025  0.09455   0.09474 2.273 0.0231
   sexm 0.07673   1.07975  0.09465   0.09519 0.806 0.4202

1:3         coef exp(coef) se(coef) robust se      z      p
   trtB -0.5214    0.5937   0.4701    0.4665 -1.118 0.2636
   sexm  1.1557    3.1763   0.4740    0.4726  2.445 0.0145

…

 States:  1= (s0), 2= CR, 3= relapse, 4= death

Likelihood ratio test=30.83  on 12 df, p=0.002093
n= 1326, number of events= 1000
```

Tidy the fit

```
> ans <- broom::tidy(cfit)
> ans$var <- substr(ans$term,1,4)
> ans$transition <- substr(ans$term,6,8)
```

transition	var	estimate	robust.se	p.value
1:2	trtB	0.21530899	0.09474336	0.02305334
1:2	sexm	0.07673060	0.09519089	0.42020202
1:3	trtB	-0.52140969	0.46645122	0.26364280
1:3	sexm	1.15571274	0.47260361	0.01446849
2:3	trtB	-0.16244028	0.14179361	0.25195624
2:3	sexm	-0.23386320	0.14228547	0.10025478
1:4	trtB	-0.09678877	0.19708123	0.62334793
1:4	sexm	0.35235819	0.19756517	0.07450436
2:4	trtB	-0.65278621	0.29374776	0.02626521
2:4	sexm	0.09211181	0.28844140	0.74946624

MULTISTATE MODEL WITH CONSTRAINTS

- All transitions to death have same coefficient

```
> cfit2 <- coxph(list(Surv(tstart, tstop, event) ~
                          trt + sex,
                      1:4 + 2:4 + 3:4 ~
                          trt + sex / common),
                 data = m1, id = id)

> round(coef(cfit2, matrix=TRUE), 3)

         1:2    1:3    2:3    1:4    2:4    3:4
trtB 0.215 -0.521 -0.162 -0.294 -0.294 -0.294
sexm 0.077  1.156 -0.234  0.228  0.228  0.228
attr(,"states")
[1] "(s0)"    "CR"    "relapse" "death"
```

CHECK PH ASSUMPTION

- A score test for proportional hazards (PH) is obtained using the `cox.zph` function, just as in a standard `coxph` model.

```
> cox.zph(cfit2)

          chisq df     p
trt_1:2  0.8465  1 0.358
sex_1:2  2.3974  1 0.122
trt_1:3  0.0666  1 0.796
sex_1:3  0.8907  1 0.345
trt_2:3  3.4882  1 0.062
sex_2:3  1.4994  1 0.221
trt_1:4  0.3161  1 0.574
sex_1:4  1.6168  1 0.204
GLOBAL  11.1414  8 0.194
```

 |

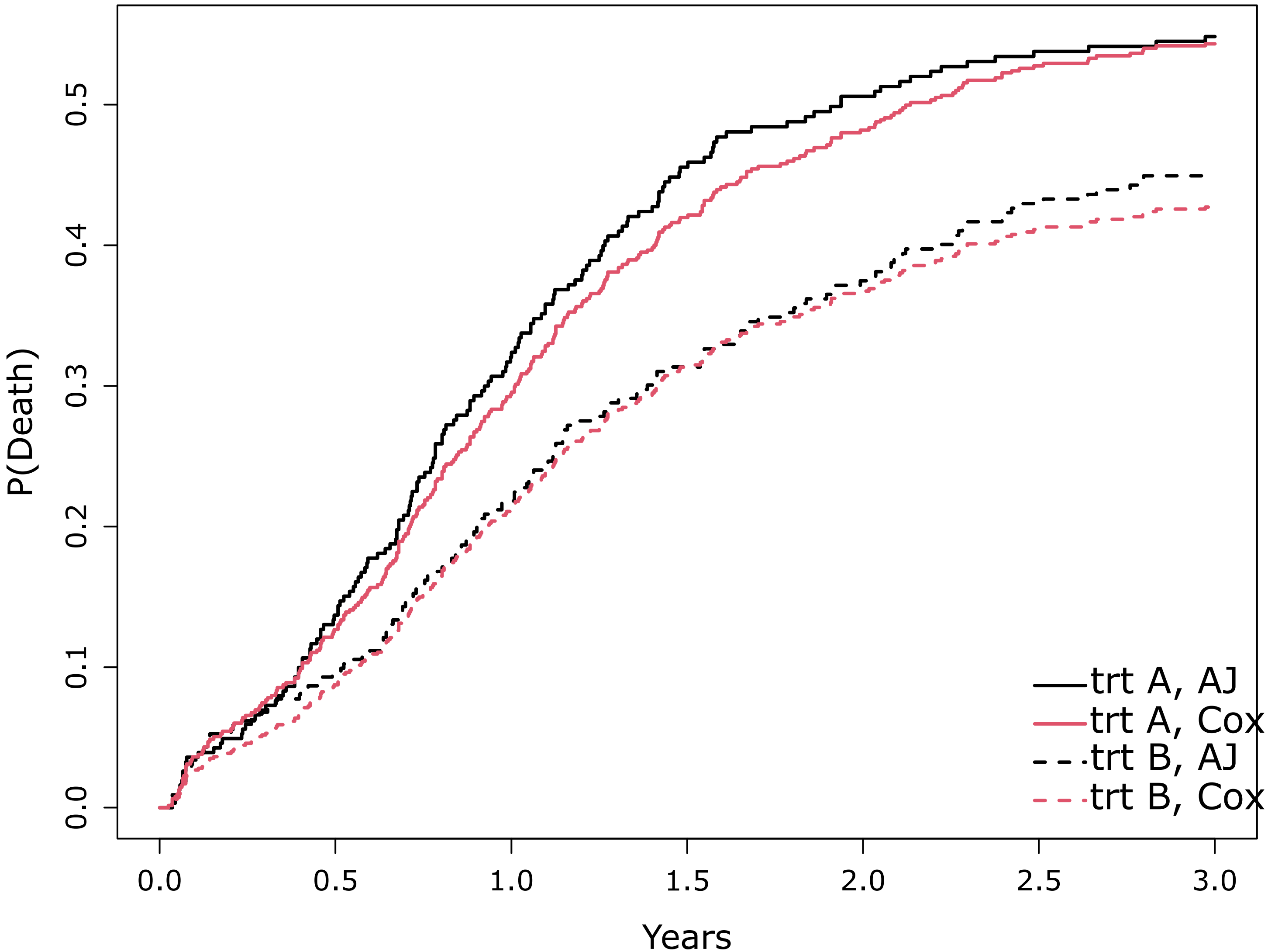
PREDICTED CURVES

```
> dummy <- data.frame(trt=c("A", "B"), sex = "f")
> csurv2 <- survfit(cfit2, newdata=dummy)
> dim(csurv2)
  data states
     2      4
```

```
> plot(fit1[,4], lty=1:2, lwd=1, xmax= 3*365,
     xscale=365, xlab="Years", ylab="P(Death)")
```

```
> lines(csurv2[,4], lty=1:2, lwd=2, col=2)
> legend("bottomright", c("trt A, AJ", "trt A, Cox",
                          "trt B, AJ", "trt B, Cox"),
      lty=c(1,1,2,2), lwd=c(1,2,1,2),
      col=c(1,2,1,2), bty='n')
```


PREDICTED CURVES



OTHER PACKAGES

- {mstate}
  - Most functionality is now available in {survival}
  - Vignette shows comparison.
  https://github.com/therneau/survival/blob/master/vignette2/tutorial.pdf

- {msm}
  - Fit multistate models where the transition time between states is not observed (such as information only detected at medical appointments)

CONCLUSION

- Lots of new functionality within the survival package
  - pseudo(), yates(), rttright(), Surv2(), concordance()
- Creating dataset the most challenging part of multistate models
- Tools are tested, easy to use

