

# Multi-state models

Terry Therneau

Department of Health Science Research  
Mayo Clinic

Nov 2017



## Survival Data

## Aalen-Johansen Curves

## Overview

## Competing risk

## Fine-Gray model

## Multiple disease states

## Population averages

## Time-dependent covariates



# Context

- ▶ I am a statistician working in medical research.
- ▶ Mayo is a tertiary care center
- ▶ Most of the question I work with are “time until ...”
  - ▶ death due to advanced cancer
  - ▶ recurrent episodes in Crohn’s disease
  - ▶ waiting time until organ transplant
  - ▶ ...

# Censoring

Key issue: it is time to do the analysis, and not every subject has yet had an event.

This is most often encoded as a pair of variables using 0/1 for the status where 1= complete observation and 0= censored.

```

> library(survival)
> test <- data.frame(time=    c(9, 3,1,1,6,6,8),
                      status=c(1,NA,1,0,1,1,0),
                      x=      c(0, 2,1,1,1,0,0))

> test
  time status x
1    9      1 0
2    3     NA 2
3    1      1 1
4    1      0 1
5    6      1 1
6    6      1 0
7    8      0 0

> #
> Surv(test$time, test$status)
[1] 9  3? 1  1+ 6  6  8+

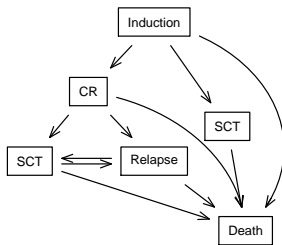
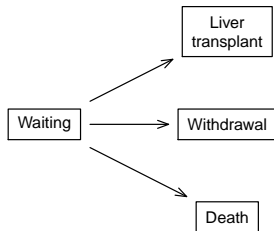
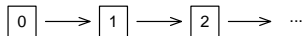
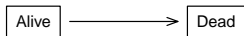
```

# Methods

- ▶ “time” as incomplete data
  - ▶  $(t, \delta)$  and covariates  $X$
  - ▶ The traditional viewpoint
  - ▶ Won't be seen again.

# Methods

- ▶ “time” as incomplete data
  - ▶  $(t, \delta)$  and covariates  $X$
  - ▶ The traditional viewpoint
  - ▶ Won't be seen again.
- ▶ Counting process view
  - ▶ Subjects go from state to state
  - ▶ Some may have many transitions
  - ▶ Some may have zero
  - ▶ There is no “incomplete” data
  - ▶ Much easier to think about multiple events





# Quantities

- ▶ 1. Event rates (arrows):  $\lambda_{jk}$
- ▶ 2. Probability in state:  $p(t) = (p_1, p_2, \dots, p_k)(t)$
- ▶ 3.  $E(\text{time in state})$
- ▶ 4.  $\Pr(\text{ever visit a state})$  or lifetime risk
- ▶ 5. Visit times for a state
- ▶ Number 1 is not enough

# Quantities

- ▶ 1. Event rates (arrows):  $\lambda_{jk}$
- ▶ 2. Probability in state:  $p(t) = (p_1, p_2, \dots, p_k)(t)$
- ▶ 3.  $E(\text{time in state})$
- ▶ 4.  $\Pr(\text{ever visit a state})$  or lifetime risk
- ▶ 5. Visit times for a state
- ▶ Number 1 is not enough
- ▶ Statisticians in the field tend to flip back and forth between 1 and 2, which can confuse onlookers.

# Event rates

- ▶ Simple rate  $r = \sum d_i / \sum t_i$   
 $P(T > t) = \exp(-rt)$
- ▶ labeled as  $r(t)$ ,  $h(t)$ ,  $\lambda(t)$ ,  $\alpha(t)$
- ▶ Underpin
  - ▶ Kaplan-Meier curves
  - ▶ Proportional hazards (Cox) model
  - ▶ Log-rank test

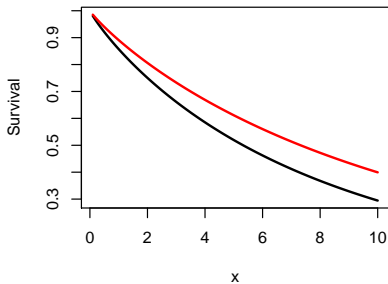
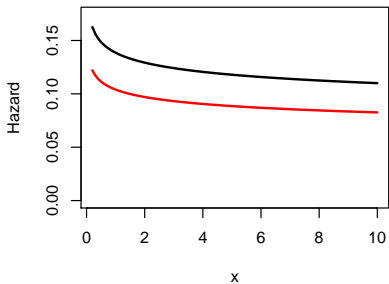
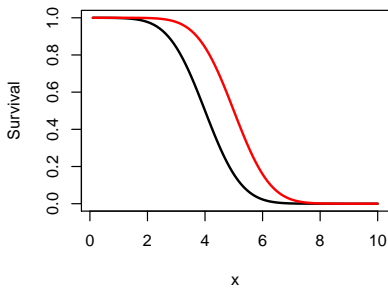
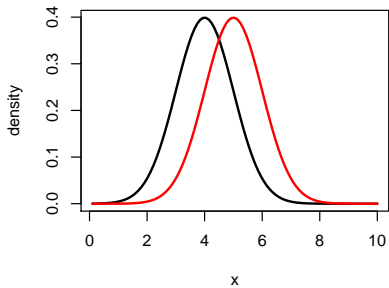
# Event rates

- ▶ Simple rate  $r = \sum d_i / \sum t_i$   
 $P(T > t) = \exp(-rt)$
- ▶ labeled as  $r(t)$ ,  $h(t)$ ,  $\lambda(t)$ ,  $\alpha(t)$
- ▶ Underpin
  - ▶ Kaplan-Meier curves
  - ▶ Proportional hazards (Cox) model
  - ▶ Log-rank test
- ▶ Martingale theory gives a formal underpinning.
  - ▶  $(N(t), Y(t))$  and  $X(t)$  or  $Z(t)$
  - ▶  $N_{ij}(t)$  = number of events so far, subject  $i$ , event type  $j$   
 $Y_{ij}(t) = 1$  if subject  $i$  is at risk for event type  $j$  at time  $t$

# Graunt's Life Table (1662)

Age Interval	Proportion Deaths in Interval	Proportion Surviving until start of Interval
0-6	0.36	1.00
7-16	0.24	0.64
17-26	0.15	0.40
27-36	0.09	0.25
37-46	0.06	0.16
47-56	0.04	0.10
57-66	0.03	0.06
67-76	0.02	0.03
77-86	0.01	0.01



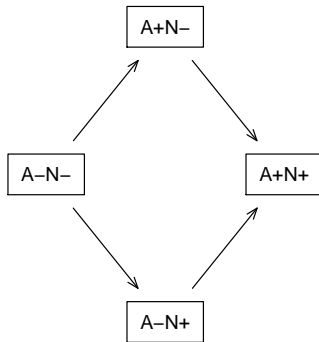
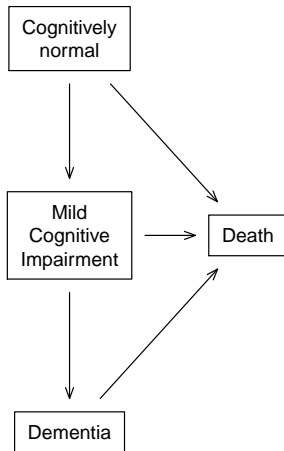








# Mayo Clinic Study of Aging



## Informative censoring

- ▶ All time to event models assume *uninformative censoring*.
- ▶ You cannot cease following someone because of something that will happen in the future.
  - ▶ Look ahead: analysis of those who "comply with the treatment"
  - ▶ People who drop out because they are about to fail MDPIT trial (Oakes, JASA 1993; 88:44-49)
  - ▶ Only those who are sick respond to queries.
  - ▶ Availability of nursing home beds.





# Transplant outcome

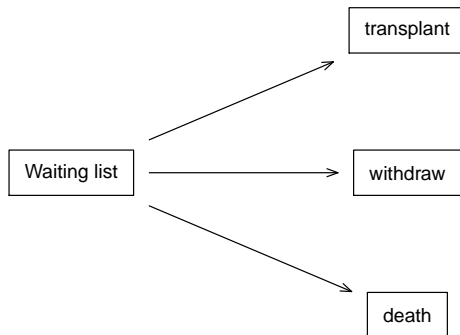
The data set that first forced me to explore multi-state methods.

- ▶ All patients added to the liver transplant waiting list during 1990-1999 at Mayo Clinic
- ▶ Corresponds to an explosion in LT programs and patients
- ▶ The median waiting time at the start is 45 days, 446 days at the end
- ▶ Question: did death on the waiting list increase?

## Note

The current liver transplant waiting list, handled by the United Network for Organ Sharing (UNOS), is managed in a much more sophisticated way than it was at the time of this data. Do not make conclusions about current conditions from this analysis.

# Competing risk





## Impact of Waiting Time

- ▶ All Mayo patients listed from Feb 1990 to Aug 1999
- ▶ 815 subjects: 636 OLT, 66 death, 37 withdraw, 76 censored

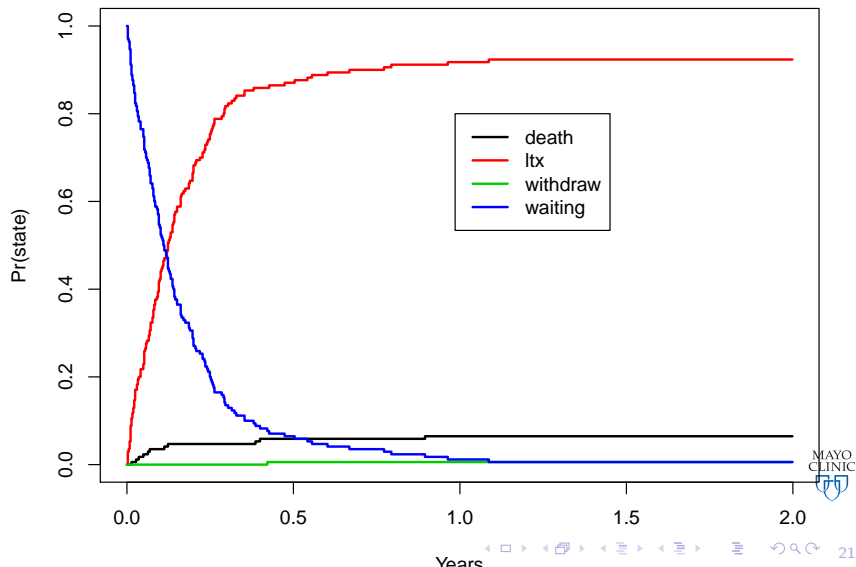


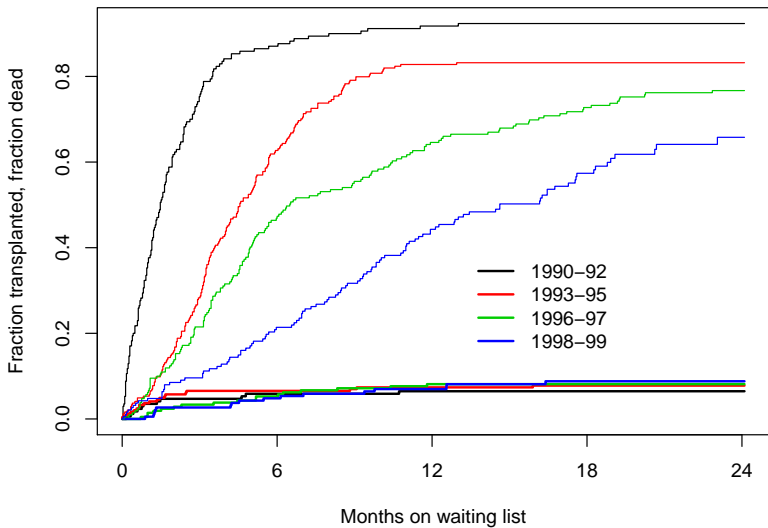
# Impact of Waiting Time

- ▶ All Mayo patients listed from Feb 1990 to Aug 1999
- ▶ 815 subjects: 636 OLT, 66 death, 37 withdraw, 76 censored
- ▶ Primary question: Did increased waiting time harm survival?  
For whom?
- ▶ transplant data set is included in R
  - ▶ age, sex, blood type, year of listing
  - ▶ follow-up time and event = (death, ltx, withdraw, censor)

# Aalen-Johansen estimate

```
> afit <- survfit(Surv(futime, event)~ 1, data=transplant)
> plot(afit, col=1:4)
```

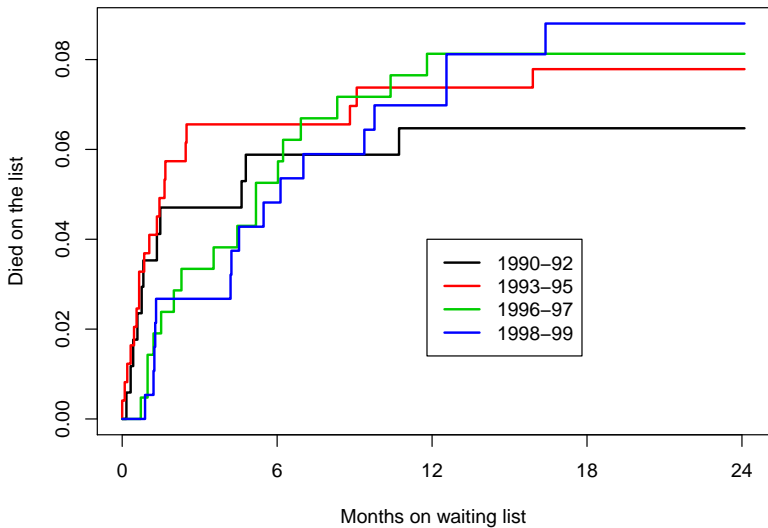


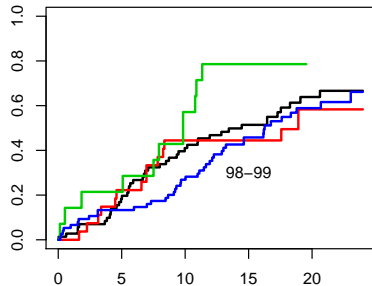
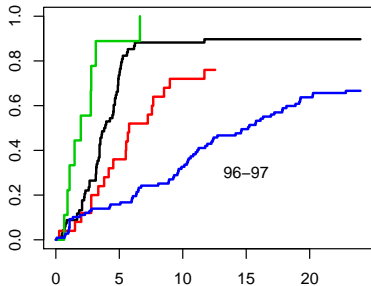
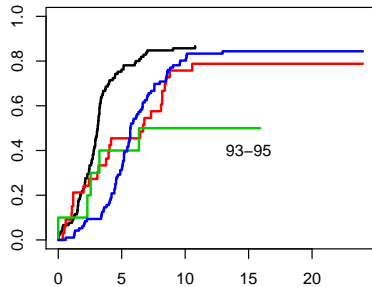
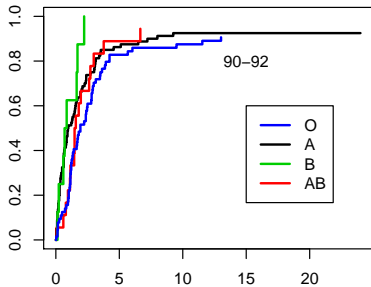


# plots

- ▶ There are 16 curves = 4 states \* 4 periods
- ▶ In R the curves can be treated like a matrix
  - ▶ `pfit[1,]` = all states, first period
  - ▶ `pfit[,1]` = death state, all periods
- ▶ Useful, since a primary display challenge is artistic: how to not be overwhelmed with a spaghetti plot.

```
> pfit <- survfit(Surv(futime, event) ~ period, transplant)
> plot(pfit[,1:2], ...)
```



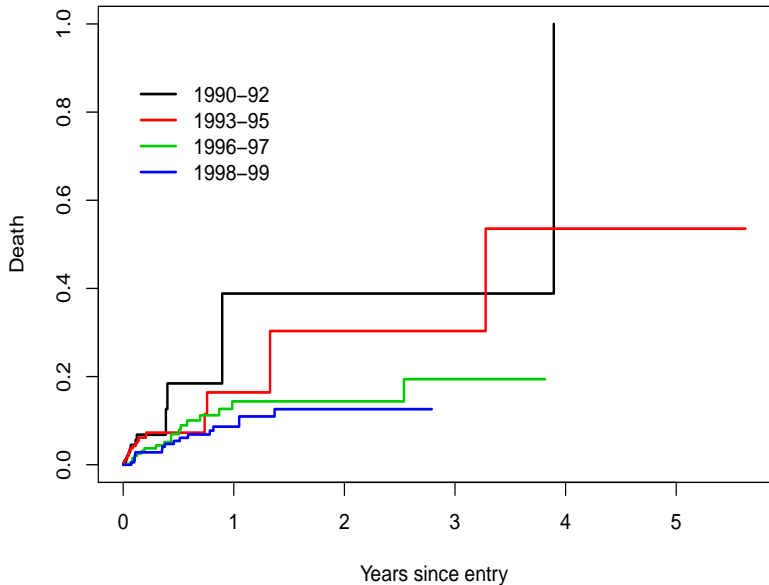


## Wrong approach

Fit a survival model of time to death, treating transplant and withdrawal as censored.

```
> badfit <- survfit(Surv(futime, event=="death") ~ period,
                    data= transplant)
> plot(badfit, fun='event', col=1:4, mark.time=F, lwd=2,
       xscale=365.25, xlab="Years since entry", ylab="Death")
> legend(.5, .9, levels(period), lty=1, lwd=2, col=1:4, bty='n')
```





## What's wrong?

- ▶ Major issues

1. An estimate of death rates for a population *where transplant and withdrawal were abolished*. That was not the question.
2. The estimate assumes uninformative censoring, i.e., those who were removed due to transplant were a random sample of those at risk for death. Rarely true, if ever.

- Unreliable estimate of an uninteresting quantity.

## What's wrong?

- ▶ Major issues
  1. An estimate of death rates for a population *where transplant and withdrawal were abolished*. That was not the question.
  2. The estimate assumes uninformative censoring, i.e., those who were removed due to transplant were a random sample of those at risk for death. Rarely true, if ever.
- ▶ Unreliable estimate of an uninteresting quantity.
- ▶ The multi-state curve estimates the fraction of subjects who will *actually experience* death before transplant.
  - ▶ Aalen-Johansen estimate replaces the Kaplan-Meier
  - ▶ an observable quantity

This class has 3 main points

- ▶ multi-state models are an important addition to your modeling toolbox
- ▶ they can yield important insights into your data
- ▶ they are easy to implement with current software

Sub-theme: interpretation can be complicated – you will be required to THINK.



## Focus on examples

- ▶ When
- ▶ How
- ▶ Usefulness

# Resources

- ▶ H. Putter, M Fiocco and R. B. Geskus, Tutorial in biostatistics: Competing risks and multi-state models. Stat in Medicine, 2007:2389-2430.
- ▶ T. Therneau, Competing Risks, R survival package.
- ▶ T. Therneau, Multi-state models, R survival package.
- ▶ J. Beyersmann, A Allignol, M. Schumacher, Competing and multistate models in R.
- ▶ R. Cook and J. Lawless, Multi-state models for the analysis of life history data. 2018

# Math: the KM estimator

Let  $t_1, t_2, \dots$  be the unique death times.

$$\lambda(t_i) = d_i/n_i \quad \text{fraction of deaths at } t_i$$

$$KM(t) = \prod_{s \leq t} [1 - \lambda(s)]$$

$$FH(t) = \prod_{s \leq t} \exp[-\lambda(s)]$$

Cox model code can use a Kaplan-Meier analog (Kalbfleish-Prentice) or a Fleming-Harrington analog (Breslow); the latter is easier to compute.





Then

$$\begin{aligned} p(t) &= p(0) \prod_{s \leq t} H(s) \\ &\approx p(0) \prod_{s < t} \exp(H(s) - \mathcal{I}) \end{aligned}$$

- ▶ the  $i, j$  element of  $P(t)$  is the probability that someone who started in state  $i$  at time 0 will be in state  $j$  at time  $t$ .
- ▶  $p(0)$  = starting distribution, usually  $(1, 0, 0, \dots)$
- ▶  $p_j(t) = \text{Pr}(\text{in state } j \text{ at time } t)$ ,  $\sum_j p_j(t) = 1$
- ▶  $p(t)$  is the Aalen-Johansen estimator
- ▶ Many (most) theory books use  $dA(t)$  (rows sum to 0) and  $I + dA(t)$  (rows sum to 1).

## Alternate view

- ▶ Treat survival as a Poisson process.
  1. Divide time into a bunch of intervals
  2. Within each interval calculate the naive rate of transition  $r(t)$  for death  
= (number making the transition) / time at risk
  3. Markov model estimate is  $p(0)[r_1 t_1][r_2 t_2] \dots$
- ▶ For a 2 state alive-dead model this yields the “actuarial” estimator.
  - ▶ Once quite common.
  - ▶ Applies directly to tabulated data (like Graunt’s).
  - ▶ Nuisance: how wide should the intervals be?





# Exercises

1. For the two state alive  $\rightarrow$  dead model, show that  $p_1(t)$  from the Aalen-Johansen estimate = Kaplan-Meier.
2. For the competing risks model show that  $p(t)$  = “cumulative incidence” estimator.
3. Explain why it took 25 years for statisticians to adopt the KM.
4. Explain why it has been over 35 years for the AJ, and it still is rare. (We are getting even slower?)

In the R survival package

```
fit <- survfit(Surv(time, status) ~ group, data=mydata)
```

- ▶ If `status` is a 0/1 or true/false variable, this gives the Kaplan-Meier + Greenwood estimate of variance
- ▶ If `status` is a categorical variable (factor), this gives the Aalen-Johansen estimate + infinitesimal jackknife estimate of variance.
- ▶ Same commands to create, print, plot, or extract values from the curves.





# AJ curves, 2 state model

- ▶  $p(t) = (\text{fraction still alive, fraction who have died})(t)$
- ▶ Some disciplines prefer to plot  $\text{prob}(\text{alive})$ ,
  - ▶ the classic survival curve that starts at 1 and goes down
  - ▶ some *insist* on it
- ▶ Some disciplines prefer to plot  $\text{prob}(\text{event})$ , which starts at 0 and goes up
- ▶ Convention and history, not science.
- ▶ Pocock and Altman, Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. Lancet 2002.

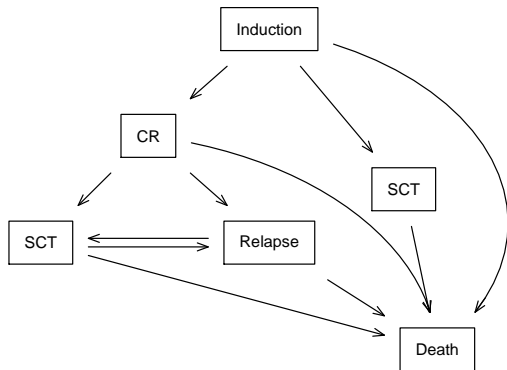
## AJ curves, multi-state

- ▶ In a multi-state model  $p(t)$  sums to 1 so one of the curves can be omitted from a plot
- ▶ Most people omit  $\text{Pr}(\text{still in the entry state})$  as it is the least interesting
- ▶ As a result all the curves go up

## AJ curves, multi-state

- ▶ In a multi-state model  $p(t)$  sums to 1 so one of the curves can be omitted from a plot
- ▶ Most people omit  $\Pr(\text{still in the entry state})$  as it is the least interesting
- ▶ As a result all the curves go up
- ▶ Some users/disciplines/journals are so bothered 'uphill!' that they suggest a plot of  $(1 - p_1(t))$ ,  $(1 - p_2(t))$ , etc.
- ▶ Imagine a plot of customers in at DFO = (fraction not in the shoe dept, fraction not in the food court, fraction not in toys, ...)



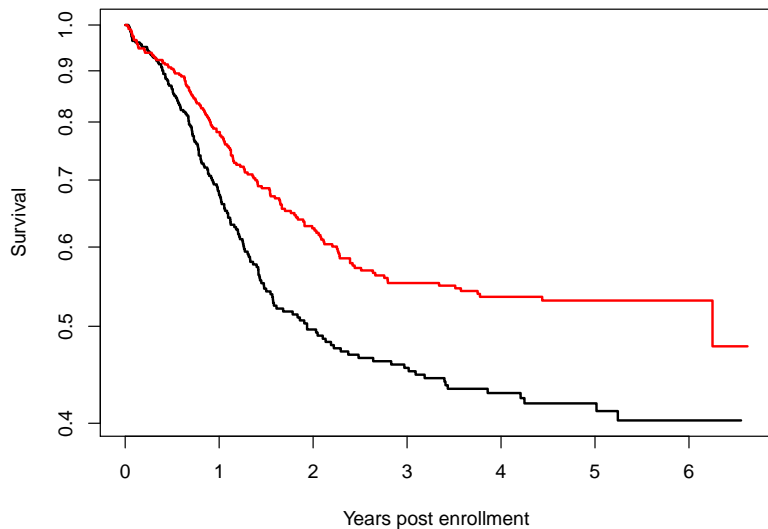


- ▶ Subjects can (and do) traverse every arrow in the diagram
- ▶ The vignette on multi-state models in the survival package explores this data set.
- ▶ [cran.r-project.org/web/packages/survival/vignettes/multi.pdf](https://cran.r-project.org/web/packages/survival/vignettes/multi.pdf)

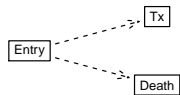
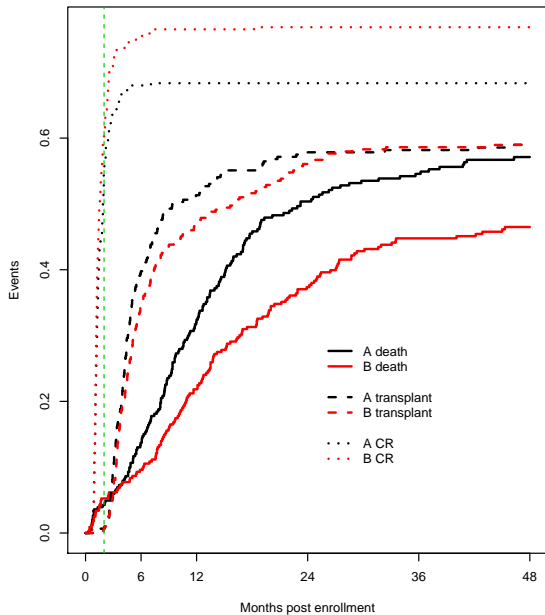
```
> myeloid[1:5,]
```

	id	trt	futime	death	txtime	crtime	rltime
1	1	B	235	1	NA	44	113
2	2	A	286	1	200	NA	NA
3	3	A	1983	0	NA	38	NA
4	4	B	2137	0	245	25	NA
5	5	B	326	1	112	56	200

# Overall survival



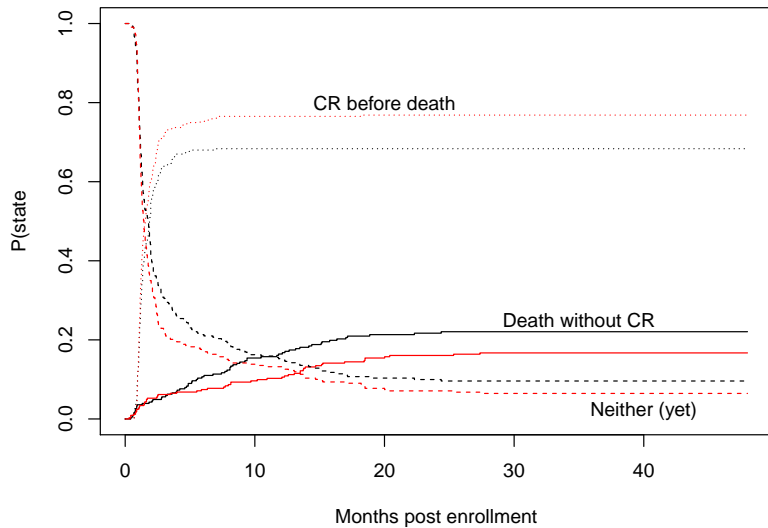


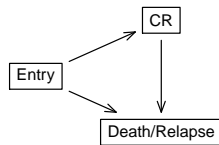
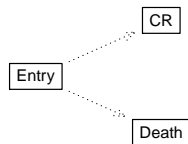
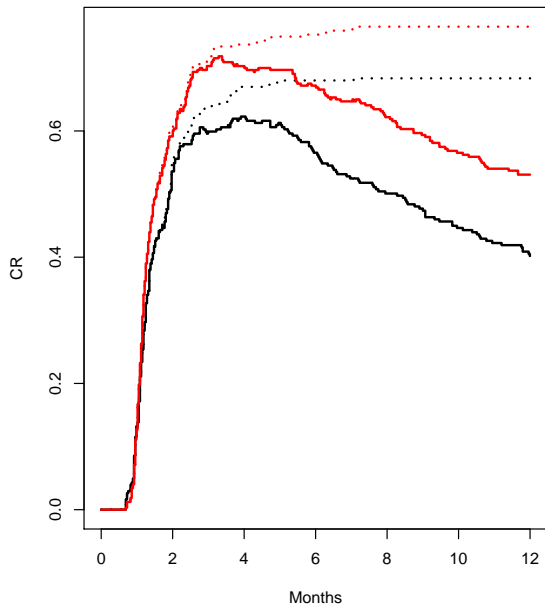


- ▶ Almost all of the CR occur by 2 months (green line)
- ▶ The additional responses for arm B happen after this



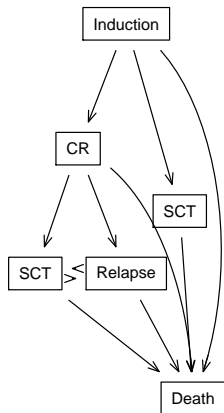
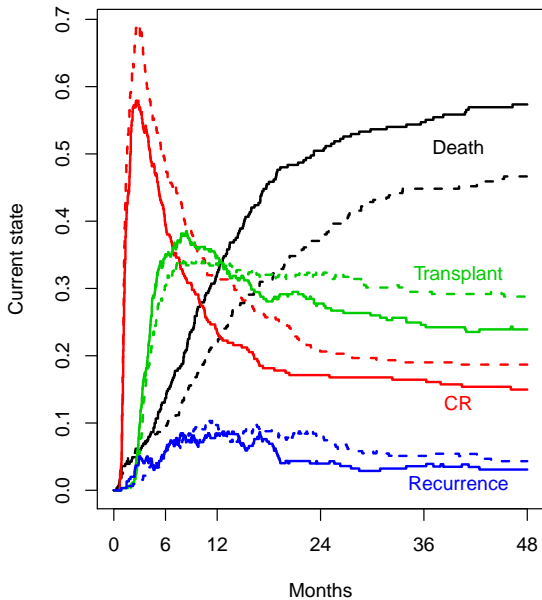






Treatment arm B has more CRs and they are more durable.

- ▶ “But I want numbers”
- ▶ Values and se at particularly follow-up times.
  - ▶ A at 6, 12, 18 months: .56(.03), .40(.03), .31(.03)
  - ▶ B at 6, 12, 18 months: .67(.03), .53(.03), .45(.03)
- ▶ Mean time in state (sojourn time)
  - ▶ *restricted* mean, up to a given time point (24)
  - ▶ A: 9.6 (0.5), B: 12.3 (0.5)
- ▶ “But I want p-values”
  - ▶  $B-A = 2.7 (.77), p < .001$



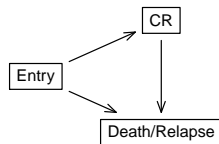
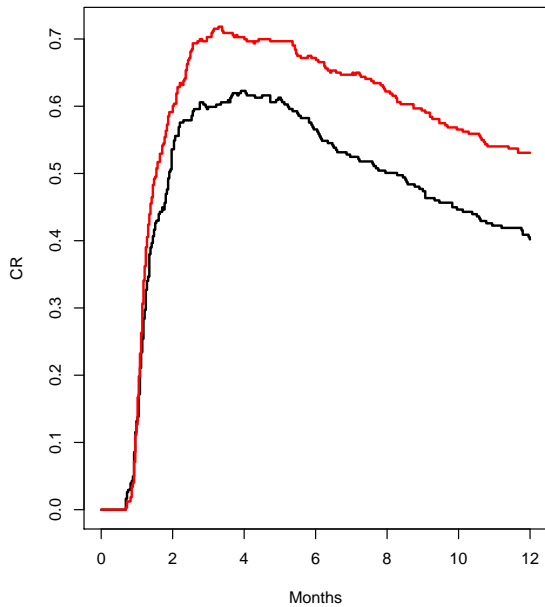




## Final figure

```
sfit4 <- survfit(Surv(tstart, tstop, event) ~ trt,  
                 data= data2, id=id)
```

	id	trt	tstart	tstop	event
1	1	B	0	44	CR
2	1	B	44	113	relapse
3	1	B	113	235	death
4	2	A	0	200	transplant
5	2	A	200	286	death
6	3	A	0	38	CR
7	3	A	38	1983	censor
8	4	B	0	25	CR
9	4	B	25	245	transplant
10	4	B	245	2137	censor





```
> sfit4 <- survfit(Surv(tstart, tstop, event) ~ trt, data2, id=id)
>
> sfit4$transitions
```

from	to			
	death	CR	transplant	relapse
death	0	0	0	0
CR	17	0	159	168
transplant	149	11	0	45
relapse	99	0	99	0
	55	443	106	13



## Warnings

Whenever the data/fit can have multiple transitions for one person

- ▶ Remember the id statement
- ▶ The program has no way of knowing which rows go together without it, and instead will assume subjects have delayed entry
- ▶ A wrong answer to the wrong question. (Variance too).





## Sequential Events







# Stacked data sets

- ▶ Diabetic retinopathy
  - ▶ 2n observations
  - ▶ Data set for the right eye, status of 0/1
  - ▶ Data set for the left eye
- ▶ Parallel failures after UCDA
  - ▶ 7 endpoints
  - ▶ 7n observations



# Sequential events

- ▶ Single stratum or multiple strata?
  - ▶ Does the baseline risk reset to a new level after each event?
  - ▶ CGD data set: no
  - ▶ Repeat cardiac events: maybe
- ▶ strata by covariate interactions?
- ▶ time scale: age, time since enrollment, time since last event,  
...



cgd

	id	treat	age	tstart	tstop	status
1	1	rIFN-g	12	0	219	1
2	1	rIFN-g	12	219	373	1
3	1	rIFN-g	12	373	414	0
4	2	placebo	15	0	8	1
5	2	placebo	15	8	26	1
6	2	placebo	15	26	152	1
7	2	placebo	15	152	241	1
8	2	placebo	15	241	249	1
9	2	placebo	15	249	322	1
10	2	placebo	15	322	350	1
11	2	placebo	15	350	439	0
12	3	rIFN-g	19	0	382	0
13	4	rIFN-g	12	0	388	0

Call:

```
coxph(formula = Surv(tstart, tstop, status) ~ treat + age +  
      cluster(id), data = cgd)
```



# Hidden covariates

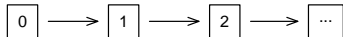
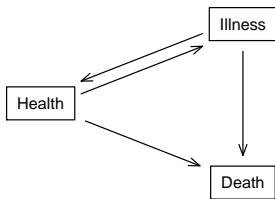
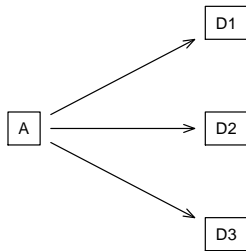
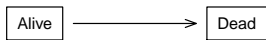
- ▶ Assume an important covariate  $Z$  is not in the model
- ▶ Single event model
  - ▶  $\beta$  biased towards zero
  - ▶ amount is proportional to  $se(\gamma Z)$
- ▶ Multiple event model
  - ▶ stratify by number of prior events, or
  - ▶ add number prior events as a covariate
  - ▶  $\beta$  is severely biased, and can actually change sign
- ▶ A random effect per subject can help

# AG simplicity

- ▶ For many studies, the coefficient(s) from an AG model often have the same interpretation as an ordinary Cox model
- ▶ higher rate  $\leftrightarrow$  shorter time to next event
- ▶ Cumulative hazard = E(number of events so far)
- ▶ Survival curve = Pr(no events at all) is more complex, but often not of interest

### 3. Competing risks





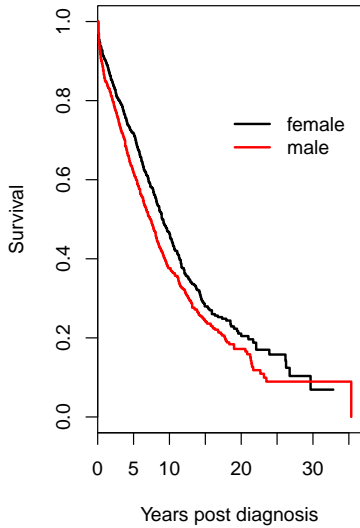
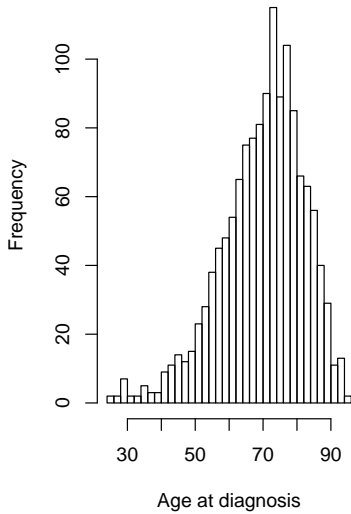
## Monoclonal Gammopathy of Undetermined Significance (MGUS)

- ▶ Subjects with a dominant clone in their plasma cell population, but without malignancy ( $\geq 2\%$  of plasma cells in the clone).
- ▶ Normally found incidentally to other tests.
- ▶ Should the patient be worried?
- ▶ About 1% per year convert to overt malignancy.
- ▶ Essentially independent of age and sex.

# Progression of MGUS

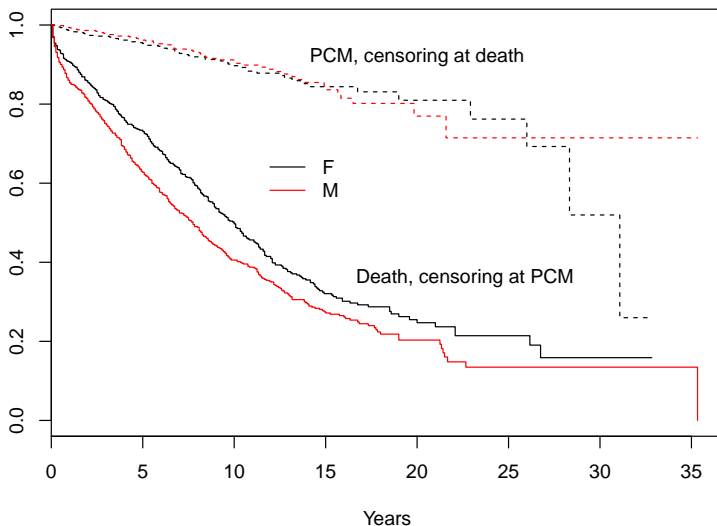
- ▶ 1384 subjects with monoclonal gammopathy of undetermined significance (MGUS)
- ▶ R. Kyle, New Engl J Med 346:564-569 (2002)
- ▶ Questions
  - ▶ Pattern of death and progression
  - ▶ Relationship to age, sex, hemoglobin, creatinine, and amount of protein in the “spike”

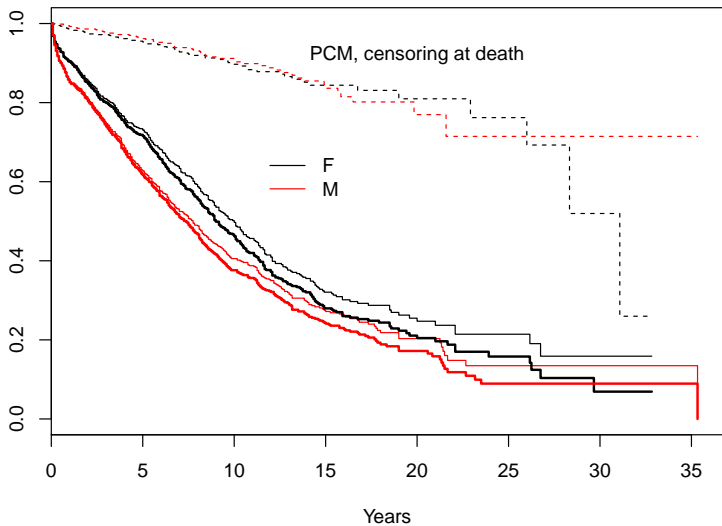






# Common analysis

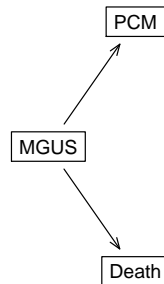
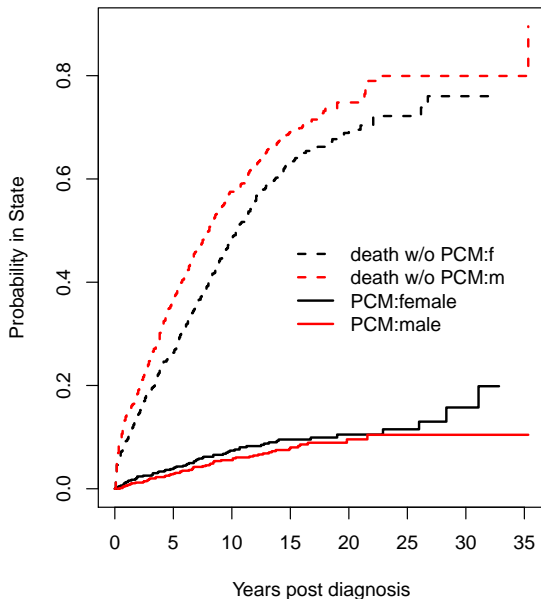




## Competing Risks

- ▶ For this illustration we are only interested in the first event for each subject.
- ▶ Formally we are treating progression to a plasma cell malignancy (PCM) as an *absorbing state*, i.e., one that subjects never exit.
- ▶ The event variable was created as a factor. The first level of the factor must be censoring, which is the status code for those whose follow-up terminated without reaching either endpoint. Codes for the remaining states can be in any order. The labels are arbitrary.
- ▶ A simple print of the `mfit2` object shows the order in which the curves will be displayed. This information was used to choose the line types and colors for the curves.
- ▶ Curves start at 0.

# Competing Risk (Aalen-Johansen)

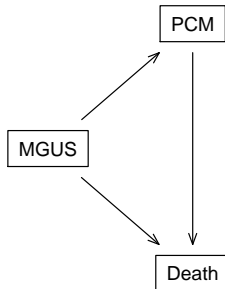
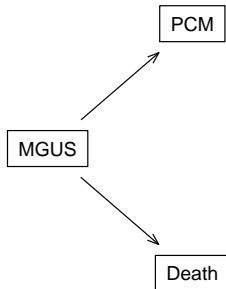


```
> mfit2 <- survfit(Surv(etime, event) ~ sex, data=mgus2)
> print(mfit2, rmean=240, scale=12)
Call: survfit(formula = Surv(etime, event) ~ sex, data = mgus2)
```

	n	nevent	rmean*
sex=F, pcm	631	59	1.323284
sex=M, pcm	753	56	1.064693
sex=F, death	631	370	8.823108
sex=M, death	753	490	10.260294
sex=F,	631	0	9.853608
sex=M,	753	0	8.675012

\*mean time in state, restricted (max time = 20 )

## Footnote: Alternate model



- ▶ Left figure: ever PCM and death without PCM
- ▶ Right figure: currently in PCM, ever dead
- ▶ Same status variable, different data set

## Raw data

	id	age	sex	ptime	pstat	futime	death
80	80	80	M	68	0	68	1
81	81	91	F	14	1	21	1
82	82	71	M	65	0	65	1
83	83	77	F	228	1	233	1

CR data: add two variables

	id	age	sex	ptime	pstat	futime	death	etime	event
80	80	80	M	68	0	68	1	68	death
81	81	91	F	14	1	21	1	14	pcm
82	82	71	M	65	0	65	1	65	death
83	83	77	F	228	1	233	1	228	pcm

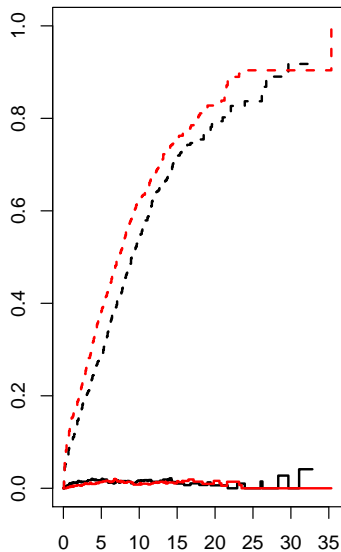
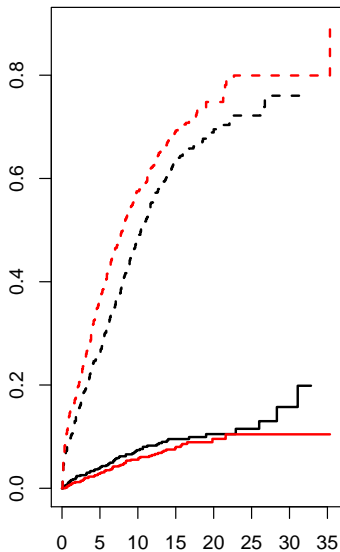
Figure 2: add lines

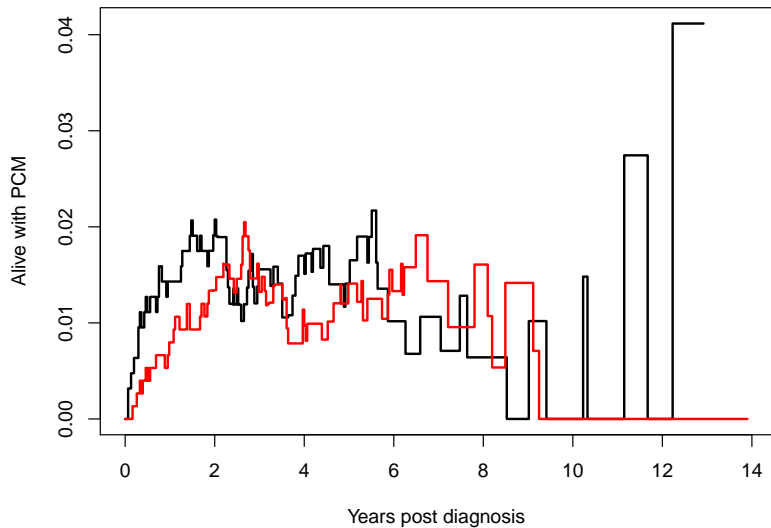
	id	age	sex	tstart	tstop	event
81	80	80	M	0	68	death
82	81	91	F	0	14	PCM
83	81	91	F	14	21	death
84	82	71	M	0	65	death
85	83	77	F	0	228	PCM
86	83	77	F	228	233	death



- ▶ The first version of the data set generated errors.
- ▶ 9 subjects have PCM and death declared at the same time.
- ▶ Treated as PCM in the CR analysis
- ▶ For the multi-state model we need to be explicit
- ▶ Push progressions back by .1 month when there is a tie

- ▶ The first version of the data set generated errors.
- ▶ 9 subjects have PCM and death declared at the same time.
- ▶ Treated as PCM in the CR analysis
- ▶ For the multi-state model we need to be explicit
- ▶ Push progressions back by .1 month when there is a tie
- ▶ Many data sets have case like this.

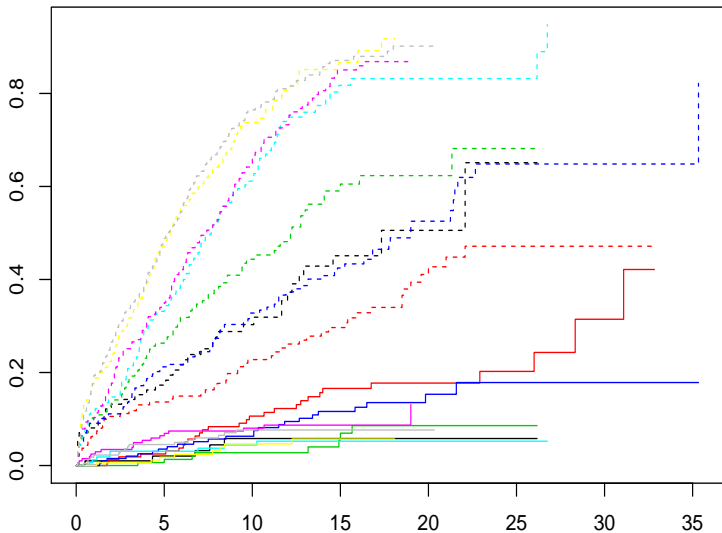




## Look at multiple factors

```
> oldage <- (mgus2$age > 70)
> bigm    <- mgus2$mspike > 1
> pfit2 <- survfit(Surv(etime, event) ~ oldage + sex +
                   bigm, data= mgus2)
>
> plot(pfit2, col=1:8, lty=rep(1:2, each=8),
       xscale=12, xlab="Years since MGUS")
```

# Tangle of yarn plot



# Hazard models

- ▶ Model the individual rates  $\lambda_{jk}(t)$  from state  $j$  to state  $k$
- ▶ Separate model for each transition
  - ▶ which covariates for each
  - ▶ which, if any, coefficients are shared
  - ▶ what time scale for each transition (baseline hazard)
  - ▶ do any transitions share a baseline
- ▶ A given arrow only depends on the starting box and the transitions
  - ▶ At risk = in the starting box
  - ▶ Event = transition of *this* type, all others are treated as censored



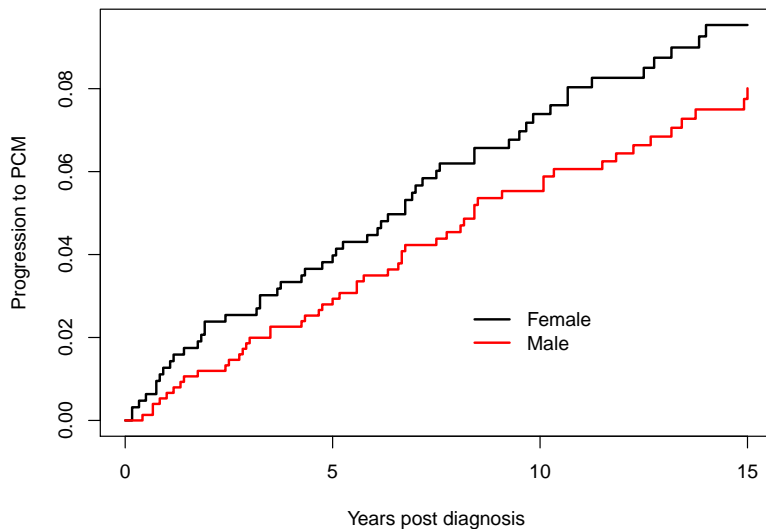


## Simple event rates

```
> pfit <- pyyears(Surv(ptime, pstat) ~ sex, scale=12*100,
                  data=mgus2)
> pfit$event/pfit$pyears
sex
      F      M
1.117354 1.016626
```

- ▶ Overall rate is 1% per year for males, 1.1% for females
- ▶ In 15 years we would expect 15–16% to progress *if there were no deaths*

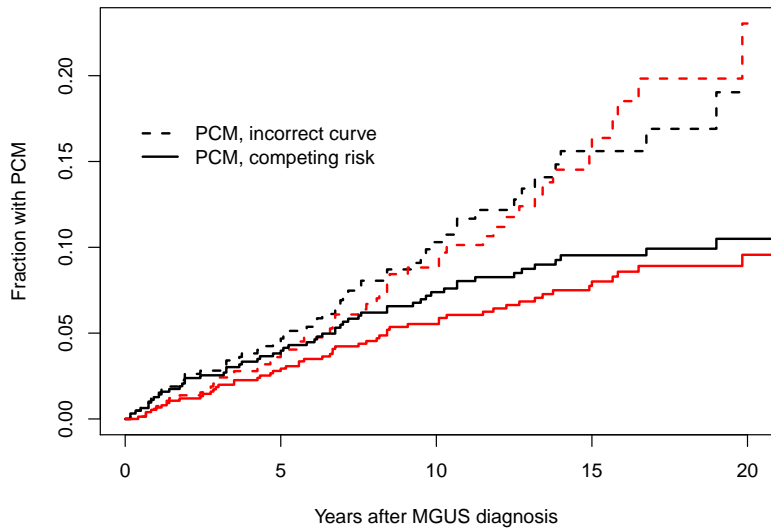
# Progression to PCM







- ▶ Underlying biology question: does gender affect the rate of progression?
- ▶ Public health question: does gender affect the lifetime risk?
- ▶ They have different answers.



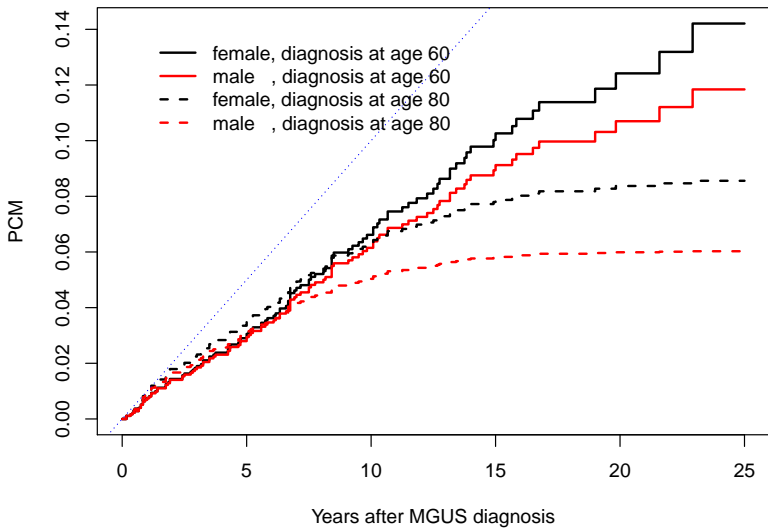






$$H = \begin{pmatrix} \lambda_{11}(t) & \lambda_{12}(t) & \lambda_{13}(t) \\ \lambda_{21}(t) & \lambda_{22}(t) & \lambda_{23}(t) \\ \lambda_{31}(t) & \lambda_{32}(t) & \lambda_{33}(t) \end{pmatrix}$$

$$= \begin{pmatrix} * & \lambda_{12}(t) & \lambda_{13}(t) \\ 0 & * & 0 \\ 0 & 0 & * \end{pmatrix}$$



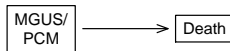
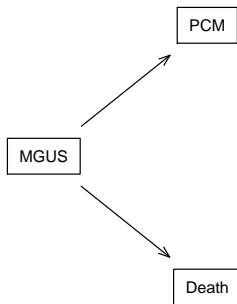














# Cox model

In a Poisson model there is a relationship between the cumulative hazard  $\lambda t$  and the CDF:

$$P(X < t) = \exp(-\lambda t)$$

An ordinary Cox model has the same relationship.

$$\lambda(t) = \lambda_0(t) \exp(X\beta)$$

$$\begin{aligned} S(t) &= \exp \left[ - \int_0^t \lambda_0(t) \exp(X\beta) \right] \\ &= p_1(t) \end{aligned}$$

where state 1 is the entry state. It is somewhat odd that there are simple expressions for the hazard of an *having* an event at  $t$  and the cumulative probability of *not having* that event by time  $t$ .

# Fine-Gray model

Rewrite as

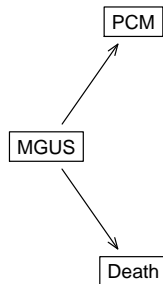
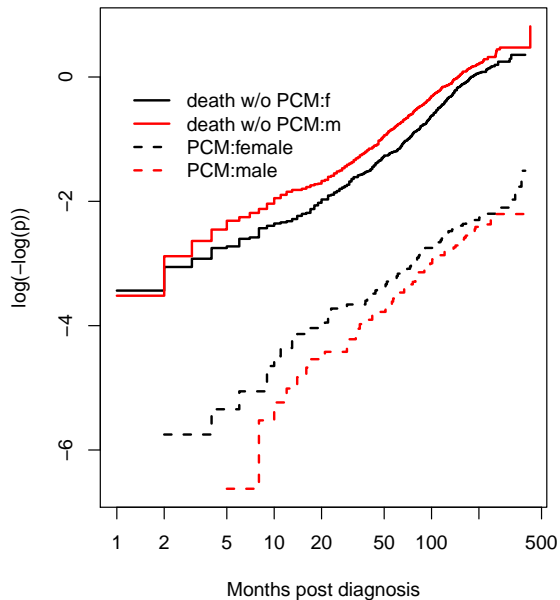
$$\begin{aligned} e^{-\Lambda(t)} &= S(t) \\ 1 - p_2(t) \\ \log(-\log(1 - p_2(t))) &= B_0(t) + \beta_1 x_1 + \dots \\ B_0(t) &= \int_0^t \beta_0(s) ds \end{aligned}$$

Looks like complimentary log-log regression





## FG works on these curves





```

> fdata1 <- finegray(Surv(etime, event) ~ id + age + sex + mspike,
                     etype = "pcm", data= mgus2)
> fdata1[1:4,]
  id age sex mspike fgstart fgstop fgstatus      fgwt
1  1  88  F   0.5      0     35          0 1.0000000
2  1  88  F   0.5     35     44          0 0.9990449
3  1  88  F   0.5     44     47          0 0.9980368
4  1  88  F   0.5     48     52          0 0.9959629
> #
> dim(mgus2)
[1] 1384  10
> dim(fdata1)
[1] 41775    8
> #
> fgfit1 <- coxph(Surv(fgstart, fgstop, fgstatus) ~ age + sex + mspike,
                  weight=fgwt, data= fdata1)

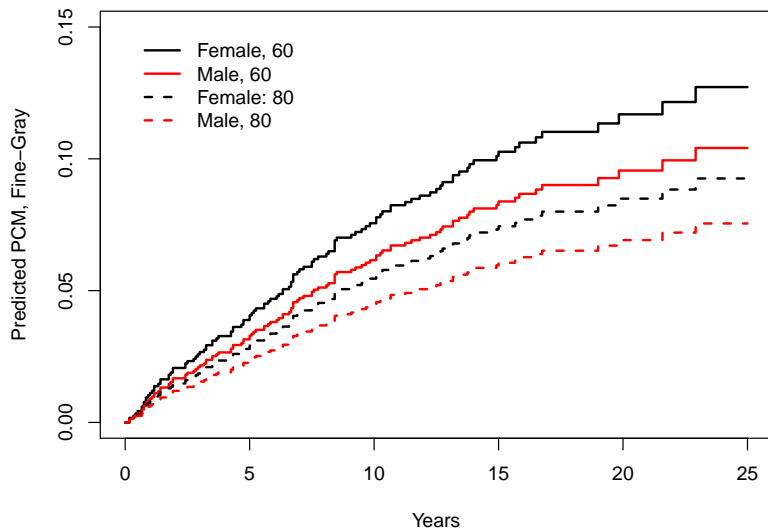
```

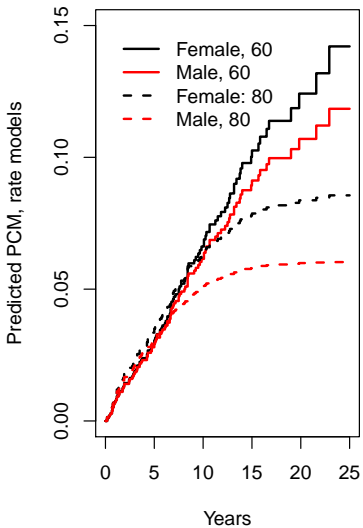
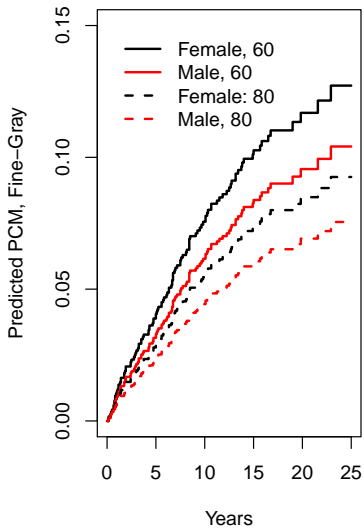
	Hazard ratio		
	age (decades)	sex (M)	serum M-spike
Cox, PCM	1.18	0.99	2.42
Fine-Gray, PCM	0.84	0.81	2.43
Cox, death	1.92	1.48	0.94
Fine-Gray, death	1.81	1.45	0.86

The raw estimates of PCM risk at 15 years were 9.5 and 8 percent for females and males, respectively, a ratio of 0.84.



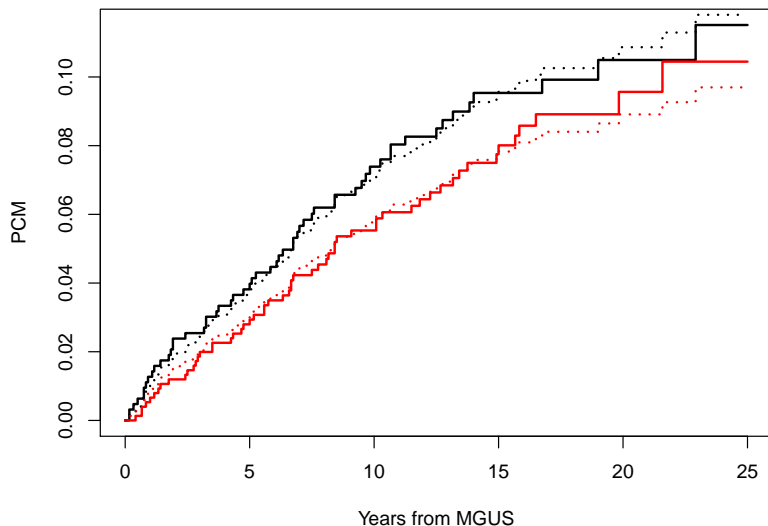
# Predicted outcome



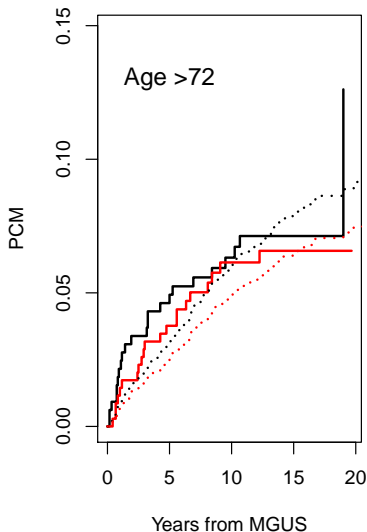
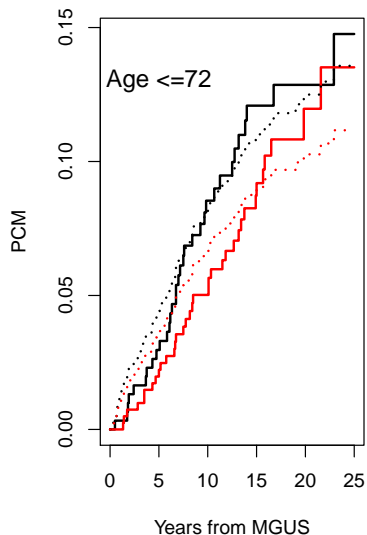




## Direct adjusted curves, total



## Direct adjusted curves, by age





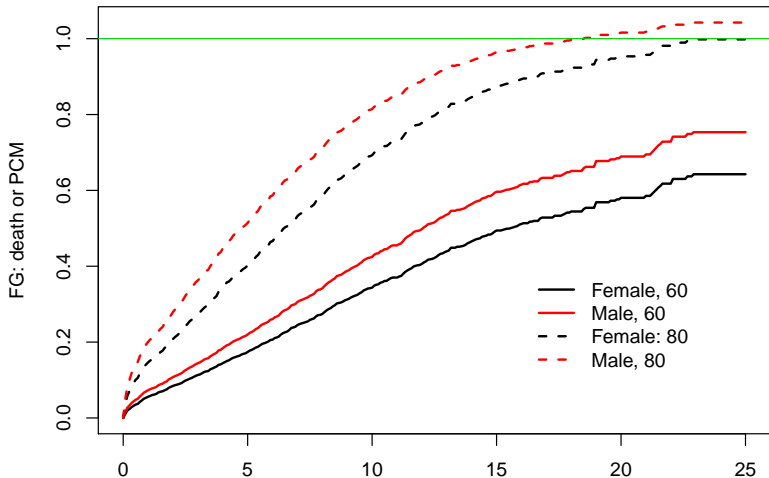


# Testing PH

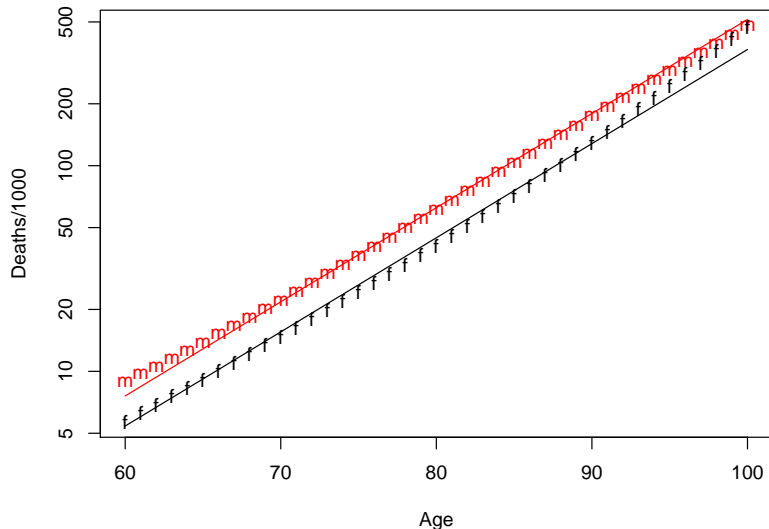
```
> cox.zph(cfit1)
              rho  chisq      p
age      -0.1736  2.3510  0.125
sexM      0.0283  0.0918  0.762
mspike   -0.0186  0.0423  0.837
GLOBAL      NA  2.5729  0.462
> #
> cox.zph(fgfit1)
              rho  chisq      p
age      -0.5306 20.9939 4.61e-06
sexM      0.0215  0.0528 8.18e-01
mspike   -0.0267  0.0788 7.79e-01
GLOBAL      NA 21.3108 9.07e-05
```

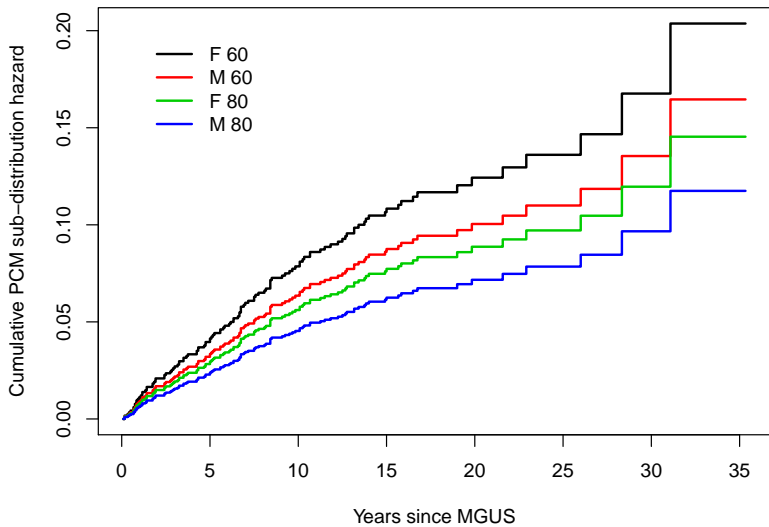


- ▶ Look at the predicted fraction who are still endpoint free.
- ▶ This is a natural part of the AJ estimate.
- ▶ The two FG estimates are separate computations; add them.



# What would an FH hazard look like?





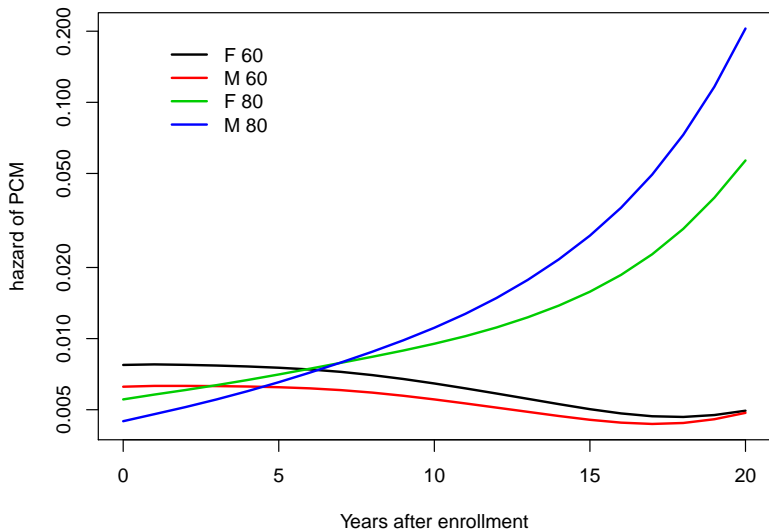
Formula 5.56 of Beyersmann, Allignol, and Schumacher

$$\alpha_1(t) = \frac{\lambda_1(t) \exp(-\Lambda_1(t) + A_2(t))}{1 - \int_1^t \lambda_1(u) \exp(-\Lambda_1(u) + A_2(u)) du}$$

- ▶  $\alpha_1, \alpha_2$  = hazard for PCM and death
- ▶  $\lambda_1$  = subdistribution hazard for PCM

$$\alpha_2(t) = \exp(-11.5 + .105(a + t) + .37m)$$

$$A_2(t) = \exp(-11.5 + .37m + .105a) (\exp(.105t) - 1) / .105$$



# Why the simple model fails for multi-state data

- ▶ Hazards are non-linear

1. In actuality there are multiple hazard operating for a subject, each with its own covariates

$$\lambda(t) = \lambda_1(t)e^{x\beta} + \lambda_2(t)e^{z\gamma} + \lambda_3(t)e^{w\psi} + \dots$$

2. PH is a model for  $\log(\lambda)$ , which does not add nicely. Multi-hazards don't collapse to a single PH equation.
3. The hazard for a heterogeneous collection of subjects is not the average of their hazards.

- ▶ Ordinary modeling issues are more acute

1. Proportional hazards rarely holds over long time periods
2. Non-linearity and interactions will often be substantial
3. Time dependent covariates are common and with particular opportunities for misuse
4. Episodic follow-up processes.
5. Informative censoring

- ▶ Model checks are imperative

## Fine-Gray

- ▶ The model often doesn't fit
  - ▶ Failure of PH on this scale
  - ▶ Particularly with long follow-up
  - ▶ Does not extend to other multi-state models
- ▶ Wrong interpretation
  - ▶ HR of .8 for sex; PCM is then interpreted as females have a higher rate, i.e., different biology.
  - ▶ We treat it as though it were a HR on one of the arrows
- ▶ Odd
  - ▶ Rate model is focused on events/(# at risk for the event)
  - ▶ FH is focused on events/(# who have not yet had the event)  
over time the denominator has more and more subjects who can never have the event
  - ▶ There is no obvious biological story that will act this way.

## Fine-Gray

- ▶ The model often doesn't fit
  - ▶ Failure of PH on this scale
  - ▶ Particularly with long follow-up
  - ▶ Does not extend to other multi-state models
- ▶ Wrong interpretation
  - ▶ HR of .8 for sex; PCM is then interpreted as females have a higher rate, i.e., different biology.
  - ▶ We treat it as though it were a HR on one of the arrows
- ▶ Odd
  - ▶ Rate model is focused on events/(# at risk for the event)
  - ▶ FH is focused on events/(# who have not yet had the event)  
over time the denominator has more and more subjects who can never have the event
  - ▶ There is no obvious biological story that will act this way.
- ▶ However
  - ▶ If the fraction with no endpoint is  $> 80\%$  the fit will often be "okay"





# Survival of the FG model?

- ▶ Has a 18 year lead on using a rate model +  $p(t)$
- ▶ It takes us 20 years for statisticians to catch on
- ▶ In the mind of many researchers FG is *the* way to deal with CR.
- ▶ FG is in R, Stata, and even SAS

Overall, SAS is about 11 years behind R and S-Plus in statistical capabilities (last year it was about 10 years behind) in my estimation.

– Frank Harrell (SAS User, 1969-1991) R-help (September 2003)

# Multiple disease states



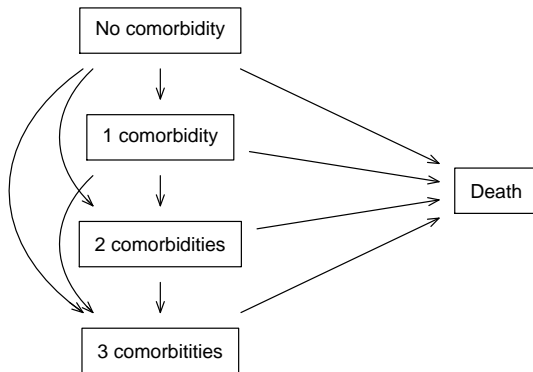
# Study

- ▶ All NAFLD diagnosis from 1997 to 2014 in Olmsted County, Minnesota.
- ▶ Utilize the Rochester Epidemiology Project
- ▶ One year delay.
- ▶ 4 controls matched on age and sex, then followed forward until the analysis date.

# Study

- ▶ All NAFLD diagnosis from 1997 to 2014 in Olmsted County, Minnesota.
- ▶ Utilize the Rochester Epidemiology Project
- ▶ One year delay.
- ▶ 4 controls matched on age and sex, then followed forward until the analysis date.
- ▶ 3864 cases of NAFLD and 14016 controls, 331 overlap.

# Target



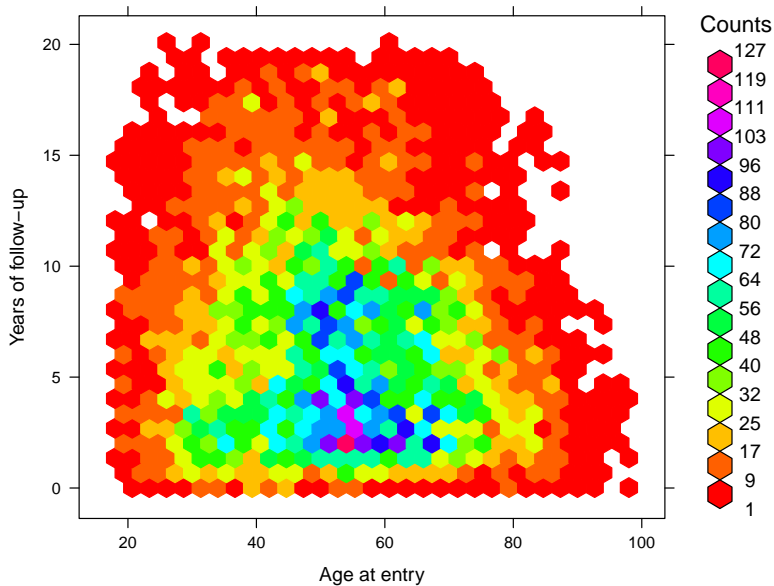
# Data

- ▶ naf1d1: One observation per subject. Baseline covariates plus follow-up time and death.
- ▶ naf1d2: Variables of id, days, test, and value. Contains selected tests and clinical observations.
- ▶ naf1d3: Variables of id, days, and event type. One observation for each outcome: occurrence of NASH, hypertension, diabetes, etc.
- ▶ To anonymize patients, all dates have been replaced with “days since first enrollment”.

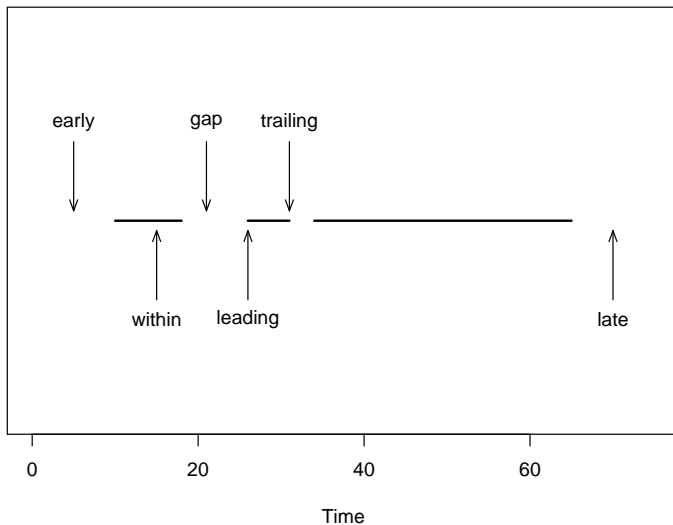


# Data

- ▶ Metabolic comorbidities are diabetes, hypertension, and dyslipidemia
- ▶ Focus on a model with 0, 1, 2, 3, of these + death
- ▶ The real work is in building and checking a data set, the fits will be easy.



# tmerge



# R code

```
> keep <- c("id", "age", "male", "bmi", "ntime")
> data1 <- tmerge(nafld1[, keep], nafld1, id,
                  death= event(futime, status))
> data1 <- tmerge(data1, subset(nafld3, event=="nafld"), id,
                  nafld = tdc(days))
> data1 <- tmerge(data1, subset(nafld3, event=="diabetes"), id,
                  diab= tdc(days), e1= event(days))
> data1 <- tmerge(data1, subset(nafld3, event=="htn"), id,
                  htn= tdc(days), e2= event(days))
> data1 <- tmerge(data1, subset(nafld3, event=="dyslipidemia"), id,
                  dyslip = tdc(days), e3= event(days))
> attr(data1, 'tcount')
```

	early	late	gap	within	boundary	leading	trailing	tied
death	0	0	0	0	0	0	17549	0
nafld	0	13	0	318	0	3533	0	0
diab	2393	0	0	1058	0	1	0	0
e1	2393	0	0	0	1058	1	0	0
htn	5022	0	0	2045	24	1	5	0
e2	5022	0	0	0	2069	1	5	0
dyslip	8663	0	0	1713	82	2	2	0
e3	8663	0	0	0	1795	2	2	0

## Four row subject

```
> rowcount <- table(data1$id)
> table(rowcount)      # pick someone with 4 rows
rowcount
      1      2      3      4      5
13501 3122  776  140   10
> subset(data1, id == 135, c(id, age, tstart, tstop, nafld,
                             htn, diab, dyslip, death))
```

	id	age	tstart	tstop	nafld	htn	diab	dyslip	death
159	135	40	0	355	1	0	0	0	0
160	135	40	355	2133	1	0	0	1	0
161	135	40	2133	3220	1	1	0	1	0
162	135	40	3220	5269	1	1	1	1	0



	id	days	event
252	135	0	nafld
253	135	355	dyslipidemia
254	135	2133	htn
255	135	2343	sleep apnea
256	135	3220	diabetes

```

> test <- tmerge(nafld1[, 1:2], nafld1, id,
                 death = event(futime, status))
> attr(test, "tcount")
      early late gap within boundary leading trailing tied
death      0      0      0      0      0      0      0      17549      0
> #
> subset(test, id==135)
      id age tstart tstop death
135 135  40      0  5269      0

```



```

> test <- tmerge(nafld1[, 1:2], nafld1, id,
                 death = event(futime, status))
> test <- tmerge(test, subset(nafld3, event=="nafld"), id,
                 nafld = tdc(days))
>
> attr(test, "tcount")
      early late gap within boundary leading trailing tied
death      0    0  0      0          0      0      17549    0
nafld      0   13  0   318          0   3533          0    0
> #
> subset(test, id==135)
      id age tstart tstop death nafld
138 135  40      0  5269      0      1

```

```

> test <- tmerge(nafld1[, 1:2], nafld1, id,
                 death = event(futime, status))
> test <- tmerge(test, subset(nafld3, event=="nafld"), id,
                 nafl = tdc(days))
> test <- tmerge(test, subset(nafld3, event=="diabetes"), id,
                 diab= tdc(days), e1= event(days))
> attr(test, "tcount")
      early late gap within boundary leading trailing tied
death      0    0  0      0          0          0    17549    0
nafl       0   13  0    318          0    3533          0    0
diab    2393    0  0   1058          0          1          0    0
e1       2393    0  0      0    1058          1          0    0
> #
> subset(test, id==135)
      id age tstart tstop death nafl diab e1
142 135  40      0  3220      0    1    0  1
143 135  40   3220  5269      0    1    1  0

```

```

> test <- tmerge(test, subset(nafl3, event=="htn"), id,
                  htn= tdc(days))
> attr(test, "tcount")
      early late gap within boundary leading trailing tied
death      0    0  0      0          0          0    17549    0
nafl       0   13  0    318          0    3533          0    0
diab    2393    0  0   1058          0          1          0    0
e1       2393    0  0      0    1058          1          0    0
htn     5022    0  0   2045          24          1          5    0
> #
> subset(test, id==135)
      id age tstart tstop death nafl diab e1 htn
155 135  40      0  2133      0    1    0  0  0
156 135  40   2133  3220      0    1    0  1  1
157 135  40   3220  5269      0    1    1  0  1

```

```
> test <- tmerge(test, subset(nafld3, event=="dyslipidemia"), id=
  lip= tdc(days), e3= event(days))
```

```
> attr(test, "tcount")
```

	early	late	gap	within	boundary	leading	trailing	tied
death	0	0	0	0	0	0	17549	0
nafl	0	13	0	318	0	3533	0	0
diab	2393	0	0	1058	0	1	0	0
e1	2393	0	0	0	1058	1	0	0
htn	5022	0	0	2045	24	1	5	0
lip	8663	0	0	1713	82	2	2	0
e3	8663	0	0	0	1795	2	2	0

```
> #
```

```
> subset(test, id==135)
```

	id	age	tstart	tstop	death	nafl	diab	e1	htn	lip	e3
159	135	40	0	355	0	1	0	0	0	0	1
160	135	40	355	2133	0	1	0	0	0	1	0
161	135	40	2133	3220	0	1	0	1	1	1	0
162	135	40	3220	5269	0	1	1	0	1	1	0



*In any sufficiently large sample, any outrageous thing is likely to happen. P. Diaconis and Mosteller, Method of studying coincidences, JASA 1989.*

- ▶ Someone *will* die on the same day as their diabetes diagnosis, have first NAFLD and first hypertension on the same day, or any number of other overlaps.
- ▶ Be prepared to think through these cases.

# Last additions

- ▶ age1, age2: age at start and end of interval
- ▶ cstate: number of metabolic conditions so far
- ▶ endpoint: censor, 1mc, 2mc, 3mc, death

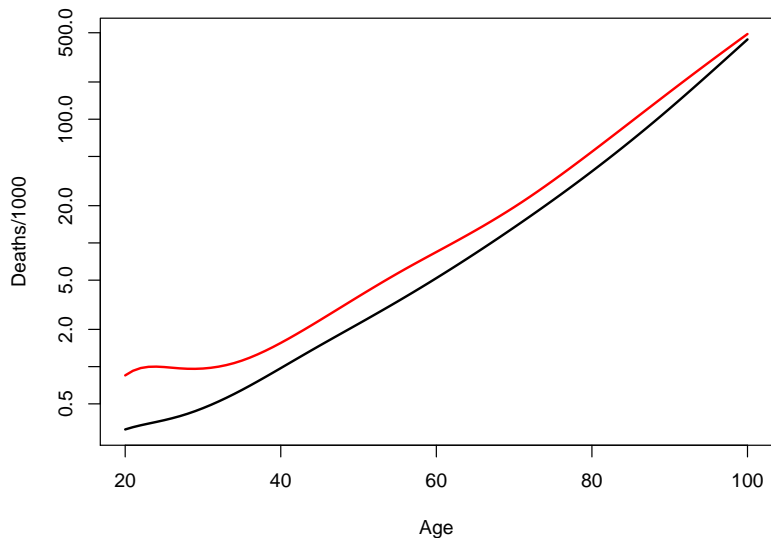
endpoint					
cstate	censored	1mc	2mc	3mc	death
0mc	5755	1829	70	4	263
1mc	4650	0	1843	28	243
2mc	3784	0	0	1048	417
3mc	2308	0	0	0	441

- ▶ Why so much time on the data?
- ▶ Print out and READ selected subjects from the final result
- ▶ If the data set is right, all that follows is easy
- ▶ If it is wrong, the answers sometimes don't show it





# 2011 Minnesota death rates



# Fits

```
> nfit1 <- coxph(Surv(age1, age2, death) ~ male + nafld,
  data=data1)
> nfit2 <- coxph(Surv(age1, age2, death) ~ male + nafld +
  as.numeric(cstate),
  data=data1)
> nfit3 <- coxph(Surv(age1, age2, death) ~ male +
  strata(cstate)/nafld, data= data1)
> nfit4a <- coxph(Surv(age1, age2, endpoint %in% c("1mc", "2mc",
  strata(male) + nafld,
  data=data1, subset= (cstate=="0mc"))
> nfit4b <- coxph(Surv(age1, age2, endpoint %in% c("2mc", "3mc")
  strata(male) + nafld,
  data=data1, subset= (cstate== "1mc"))
> nfit4c <- coxph(Surv(age1, age2, endpoint=="3mc") ~
  strata(male) + nafld,
  data=data1, subset= (cstate=="2mc"))
```

	male	sex	NAFLD	mcount
Alive -> Dead	1.4		1.6	
Alive -> Dead	1.4		1.5	1.2
0 MC -> Dead	1.4		1.9	
1 MC -> Dead	1.4		1.7	
2 MC -> Dead	1.4		1.7	
3 MC -> Dead	1.4		1.1	
0 MC -> 1+ MC	2.5		2.5	
1 MC -> 2+ MC	1.7		1.7	
2 MC -> 3 MC	1.6		1.6	

- ▶ The incremental impact of NAFLD on death decreases
- ▶ The impact on conversion to the next comorbid state decreases.
- ▶ (Allen adds these to figure 4.)

# Stacked data set

Build a transtion matrix.

from	to				
	0mc	1mc	2mc	3mc	death
0mc	-	1	2	3	7
1mc	-	-	4	5	8
2mc	-	-	-	6	9
3mc	-	-	-	-	10
death	-	-	-	-	-

```
> # data creation suppressed
> dim(msdata)
[1] 65223      23
> msdata[1:5, 1:7]
An object of class 'msdata'
```

Data:

	from	to	trans	status	id	age	male
1	1	2	1	0	1	57	0
2	1	2	1	0	2	67	0
3	1	2	1	0	3	53	1
5	1	2	1	1	5	68	1
7	1	2	1	0	6	39	0

	to							
from	0mc	1mc	2mc	3mc	death	no event	total	entering
0mc	0	1829	70	4	263	5755		7921
1mc	0	0	1843	28	243	4650		6764
2mc	0	0	0	1048	417	3784		5249
3mc	0	0	0	0	441	2308		2749
death	0	0	0	0	0	1364		1364

## All at once

```
> msdata$sgrp <- factor(with(msdata, from + 4*(to==5)))  
> allfit2 <- coxph(Surv(age1, age2, status) ~  
                    sgrp:(nafld + male) + strata(from, to)  
                    data=msdata)  
> length(coef(allfit2))  
[1] 14
```



## Other summaries

- ▶ Hazard rates between states
- ▶  $p(t)$ , state vector at time  $t$
- ▶ Mean time in state
- ▶ Number of visits to each state

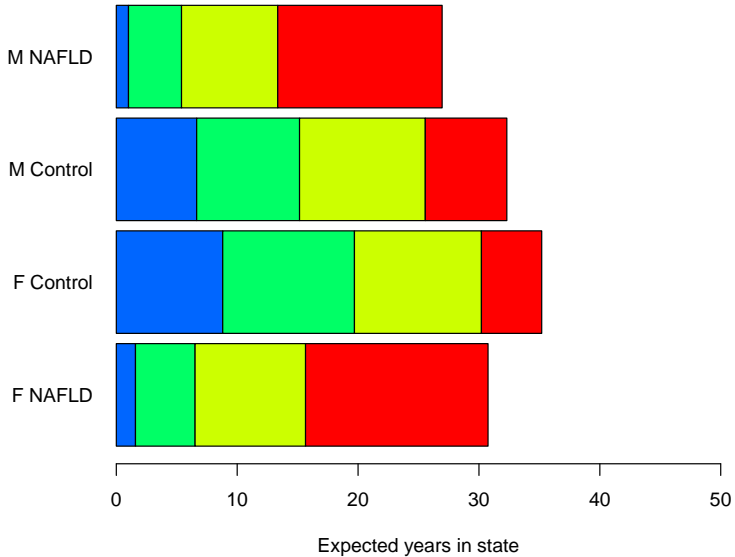


```
> multi <- survfit(Surv(age1, age2, endpoint) ~ nafld, data =  
  data, id=id, se=FALSE, start.time=0)
```

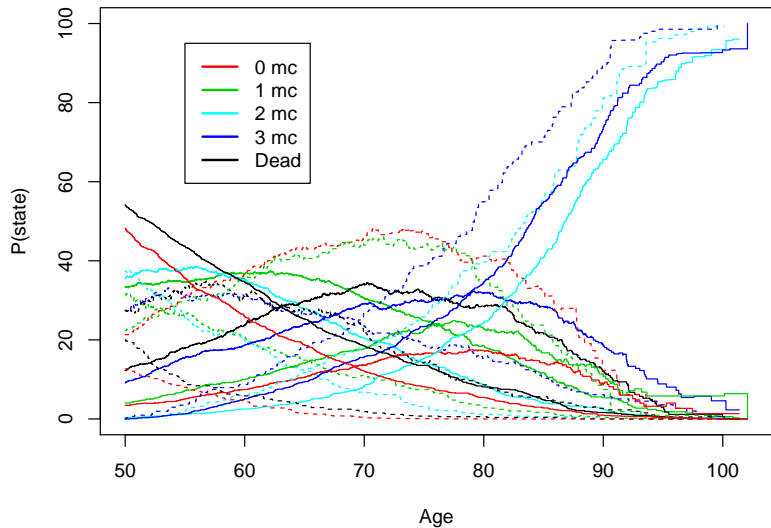
```
Error in survfitCI(X, newY, weights = casewt, id =  
id, istate = istate, : subject is in two different  
groups, id 13
```

```
> data1$ibase <- data1$naflid[match(data1$id, data1$id)]
> multi1 <- survfit(Surv(age1, age2, endpoint) ~ male + nbas
                    istate=cstate, id=id, se=FALSE, start.time
> print(multi1, digits=2, rmean=100)
Call: survfit(formula = Surv(age1, age2, endpoint) ~ male +
              data = data1, id = id, istate = cstate, se = FALSE, sta
```

	n	nevent	rmean*
male=0, nbase=0, 1mc	7188	556	10.9
male=0, nbase=1, 1mc	1791	74	4.9
male=1, nbase=0, 1mc	5935	424	8.5
male=1, nbase=1, 1mc	1492	55	4.4
male=0, nbase=0, 2mc	7188	596	10.5
male=0, nbase=1, 2mc	1791	176	9.1
male=1, nbase=0, 2mc	5935	547	10.4
male=1, nbase=1, 2mc	1492	151	8.0
male=0, nbase=0, 3mc	7188	290	5.0
male=0, nbase=1, 3mc	1791	164	15.1
male=1, nbase=0, 3mc	5935	285	6.7



	0mc	1mc	2mc	3mc	death
male=0, nbase=0	0.54	0.33	0.09	0.03	0
male=0, nbase=1	0.20	0.31	0.27	0.21	0
male=1, nbase=0	0.48	0.36	0.12	0.04	0
male=1, nbase=1	0.13	0.37	0.27	0.22	0



## Fool survfit

```
> fakeid <- data1$id + data1$nafld/2
> multi2 <- survfit(Surv(age1, age2, endpoint) ~ nafld, data = data1,
                    istate=cstate, id=fakeid, se=FALSE, start.time = 0)
> print(multi2, digits=2)
Call: survfit(formula = Surv(age1, age2, endpoint) ~ nafld, data = data1,
              id = fakeid, istate = cstate, se = FALSE, start.time = 0)
```

	n	nevent	rmean*
nafld=0, 1mc	13036	972	9.9
nafld=1, 1mc	3617	137	4.8
nafld=0, 2mc	13036	1120	10.6
nafld=1, 2mc	3617	350	8.6
nafld=0, 3mc	13036	550	5.7
nafld=1, 3mc	3617	328	14.3
nafld=0, death	13036	931	18.0
nafld=1, death	3617	369	23.0
nafld=0, 0mc	13036	0	7.8
nafld=1, 0mc	3617	0	1.3

- ▶ This fit allows someone to jump to another curve mid-stream.
- ▶ For simple alive-dead, these curves are equivalent to Simon and Makuch, Statistics in Medicine, 1984.



## Choices

1. Use only baseline values for the fit
  - ▶ Survival given that baseline
  - ▶ Subjects will evolve, and that gets built into the coefficients and the baseline hazard
  - ▶ The PH assumption may be badly strained
2. “Baseline” values at future times (landmark analysis)
3. Use time-dependent covariates for the fit
  - 3.1 Predict for fixed covariates
  - 3.2 Create a fixed covariate path (delicate)
  - 3.3 Joint fits (JM package)
  - 3.4 Use a population average
  - 3.5 Encode the path as states and refit
4. Various bad ideas

## Time in state, based on a fitted model

- ▶ Simplest method is 3.1: time-dependent fit + prediction for static covariates
  1. Fit the TD models for each state to state transition (`nfit3`, `nfit4a`, `nfit4b`, `nfit4c`)
  2. For each model, get the predicted hazard functions  $\lambda_{ij}(t, \text{control})$  and  $\lambda_{ij}(t, \text{NAFLD})$
  3. Compute Aalen-Johansen estimates using these values.
  4. Alternate: use the `mstate` package
- ▶ Better is method 1: time-fixed covariate fit + static prediction
  1. Only use the baseline `naflid` value, or create the doubled data set (full follow-up for controls that become cases)
  2. Cox model fits for each transition, robust variance if doubled
  3. Hazard estimates and AJ computation as before
- ▶ Better is method 3.3: time-dependent fit + population curves
  1. Fit the TD models
  2. Predicted hazard curves for *each covariate path* = 334 for controls, only 1 for NAFLD; overall hazard for controls is a time-weighted average.
  3. Use these to compute Aalen-Johansen values

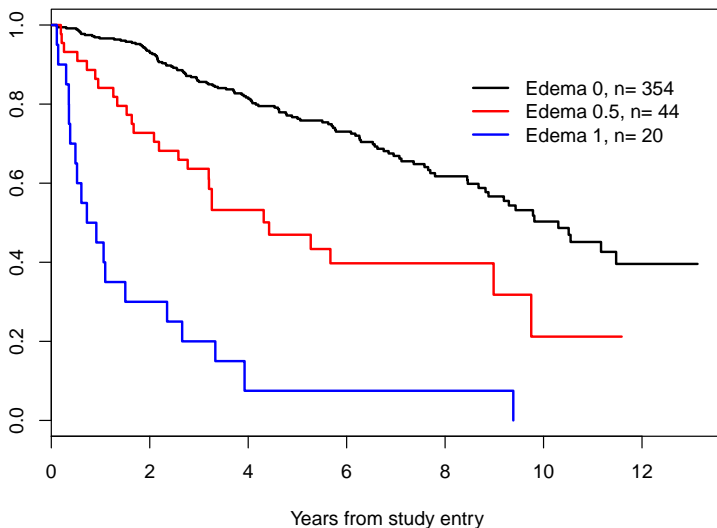


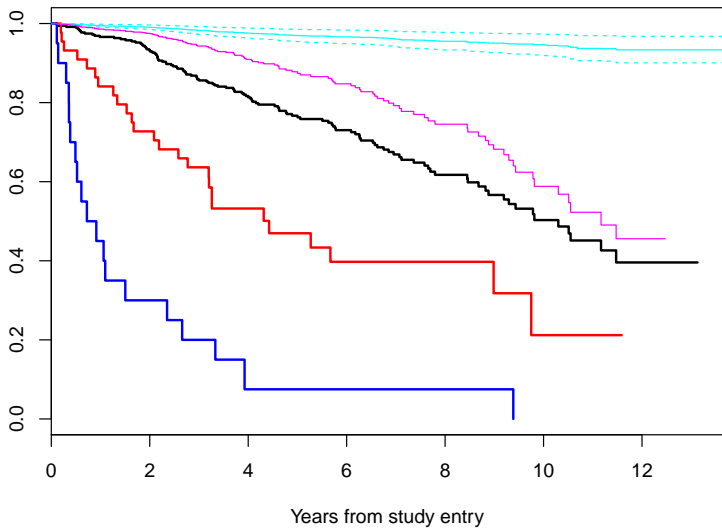


# Primary biliary cirrhosis

- ▶ Progressive autoimmune disease
- ▶ Continual inflammation slowly creates scar tissue
- ▶ Time dependent risk score fits better

# Primary biliary cirrhosis





# Using mstate



# Population averages







# Why focus on simplicity

- ▶ Terse summaries for our papers
- ▶ Too many projects on our plate
- ▶ Thoughtful simplicity: models which over-summarize are fit in order to better understand the data, but with the larger context always in mind.

“All models are wrong, the question is whether they are wrong enough to not be useful.” GEP Box

“A model is a lie that helps you see the truth.” Howard Skipper

# Why focus on simplicity

- ▶ Terse summaries for our papers
- ▶ Too many projects on our plate
- ▶ Thoughtful simplicity: models which over-summarize are fit in order to better understand the data, but with the larger context always in mind.

“All models are wrong, the question is whether they are wrong enough to not be useful.” GEP Box

“A model is a lie that helps you see the truth.” Howard Skipper

“For every complex question there is a simple and wrong solution.”

A Einstien

# Marginal estimates

- ▶ model with  $x_1, x_2, x_3, \dots$
- ▶  $PMM_{x_1=c} = E_X(\hat{y}(x)|x_1 = c)$
- ▶ Population Marginal Mean
- ▶ Idea
  - ▶ Compare treatment A to treatment B
  - ▶ Pretend we have a population of subjects = the other covariates
  - ▶ For each of those subjects we can compute the predicted response for their covariates, under treatment A and then under treatment B
  - ▶ Take an average;  $PMM_A - PMM_B$







# Old idea

- ▶  $\hat{y} = S(t)$ , population=data: direct adjusted survival
- ▶ linear model,  $\hat{y} = X\hat{\beta}$ , population=data: closely related to survey sampling estimates
- ▶ g-estimates of causal modeling — sort of
- ▶ first instinct of a statistician is to change  $Z$  to  $E(Z)$
- ▶ linear model,  $\hat{y} = X\hat{\beta}$ , population= factorial for the categoricals, data for continuous: SAS GLM type III (SGTT)

# Old idea

- ▶  $\hat{y} = S(t)$ , population=data: direct adjusted survival
- ▶ linear model,  $\hat{y} = X\hat{\beta}$ , population=data: closely related to survey sampling estimates
- ▶ g-estimates of causal modeling — sort of
- ▶ first instinct of a statistician is to change  $Z$  to  $E(Z)$
- ▶ linear model,  $\hat{y} = X\hat{\beta}$ , population= factorial for the categorical, data for continuous: SAS GLM type III (SGTT)
  1. clever and efficient formulas, but forgot what is being computed
  2. horrible documentation (document an algorithm)
  3. factorial population is rarely appropriate
  4. other SAS (and R) procedures do something different (NSTT)



# Mayo Clinic Study of Aging

- ▶ On autopsy, Alzheimer's patients have amyloid plaques and neurofibrillary tangles in their brain tissue.
- ▶ The population is aging.
- ▶ Studies
  - ▶ Alzheimer's Disease Neuroimaging Initiative (ADNI)
  - ▶ Religious Orders Study (ROS), Memory and Aging Project (MAP)
  - ▶ Mayo Clinic Study of Aging (MCSA)
    - ▶ Enroll a stratified population sample
    - ▶ Equal number of males and females, larger cohorts at older ages
    - ▶ Follow all subjects at a regular intervals
    - ▶ Replenish the cohort for drop-out and death.

## Key Measurements

- ▶ Clinical assessment
  - ▶ Cognitive tests
  - ▶ Care team



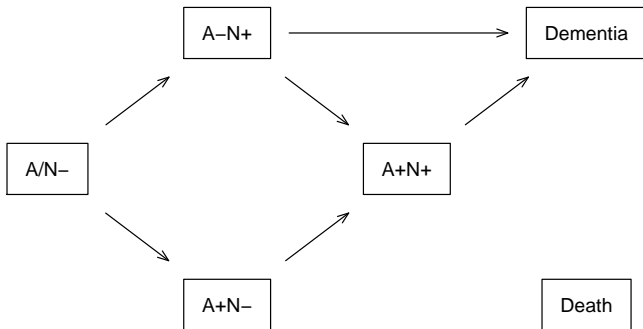
# Key Measurements

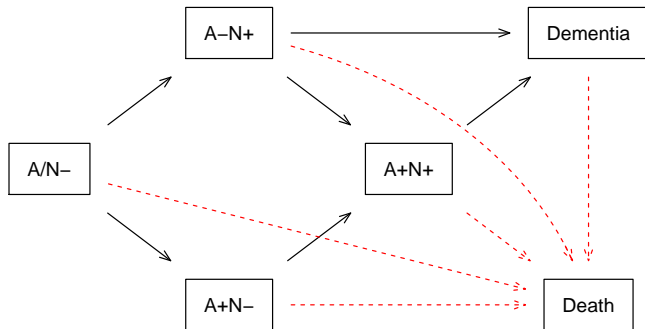
- ▶ Clinical assessment
  - ▶ Cognitive tests
  - ▶ Care team
- ▶ MRI structural scan
- ▶ Amyloid PET
- ▶ FDG PET
- ▶ Tau PET

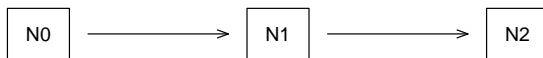
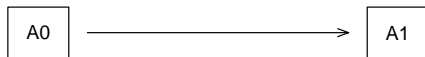












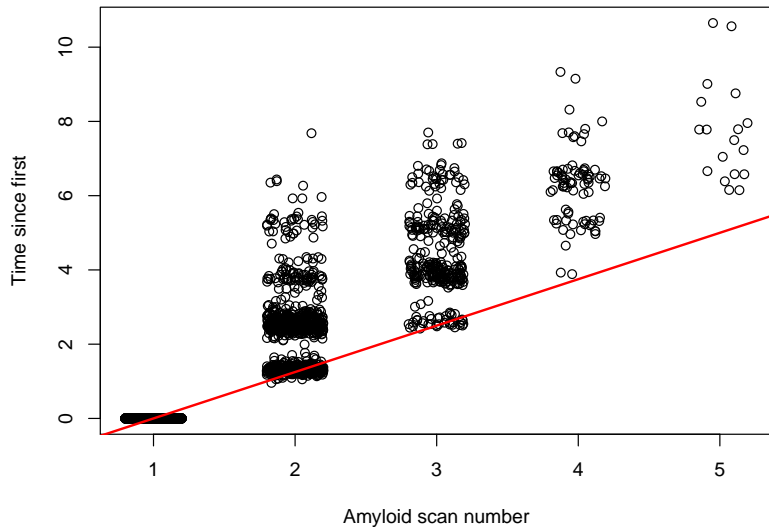
# States

- ▶ A0/A1: none/mild vs moderate/severe amyloid burden
- ▶ T0/T1: none/mild vs moderate/severe tau burden
- ▶ N0/N1/N2: increasing neurodegeneration
- ▶ 13 states (boxes)
- ▶ 32 transitions (arrows)



- ▶ 5 covariates: intercept, age, sex, APOE positivity, hypertension
- ▶ 32 transitions
- ▶ 160 potential parameters
- ▶ plus HMM parameters
- ▶ Don't get carried away!

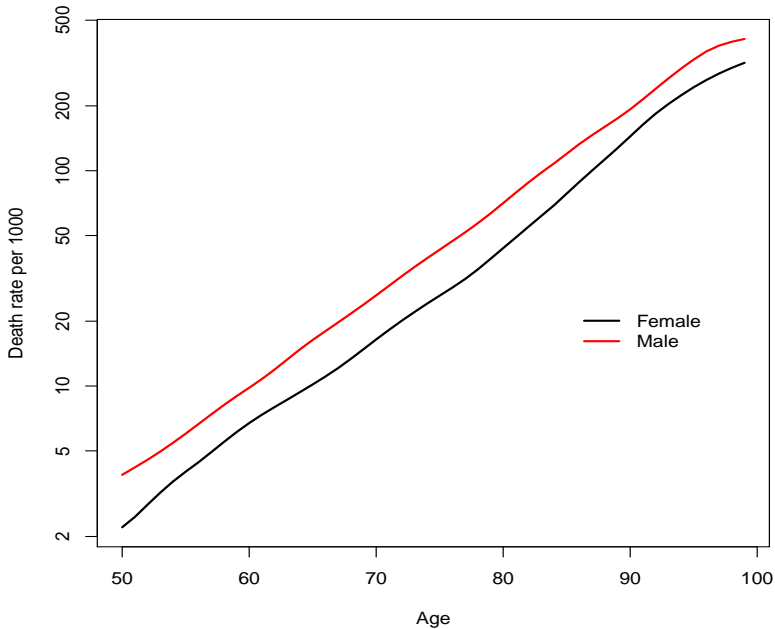
# Amyloid scan timing

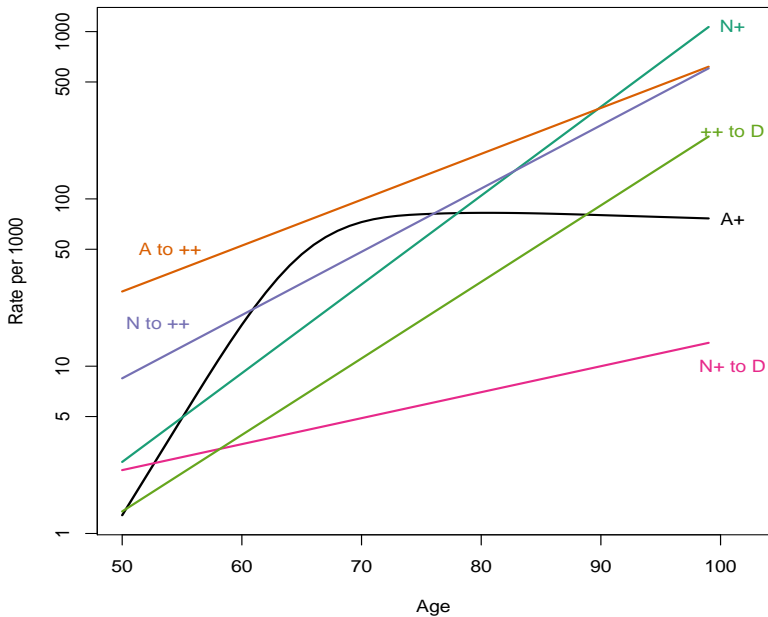


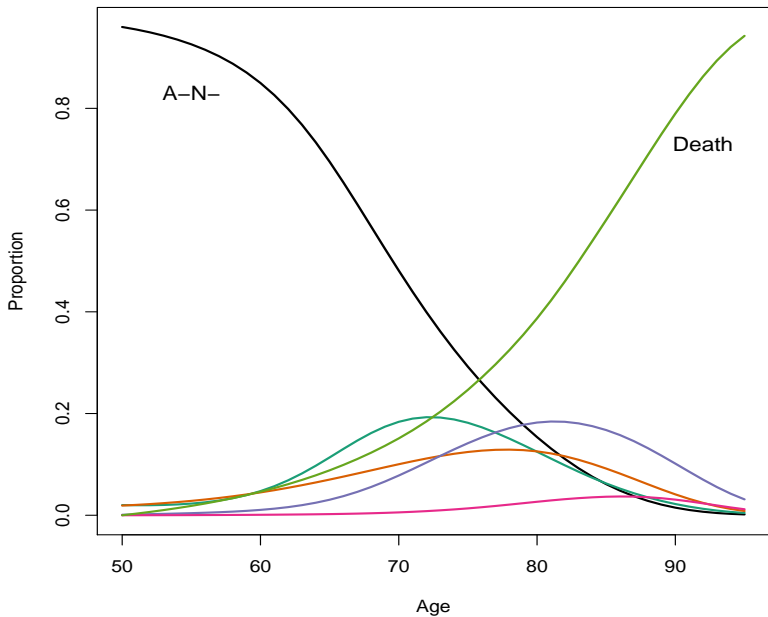


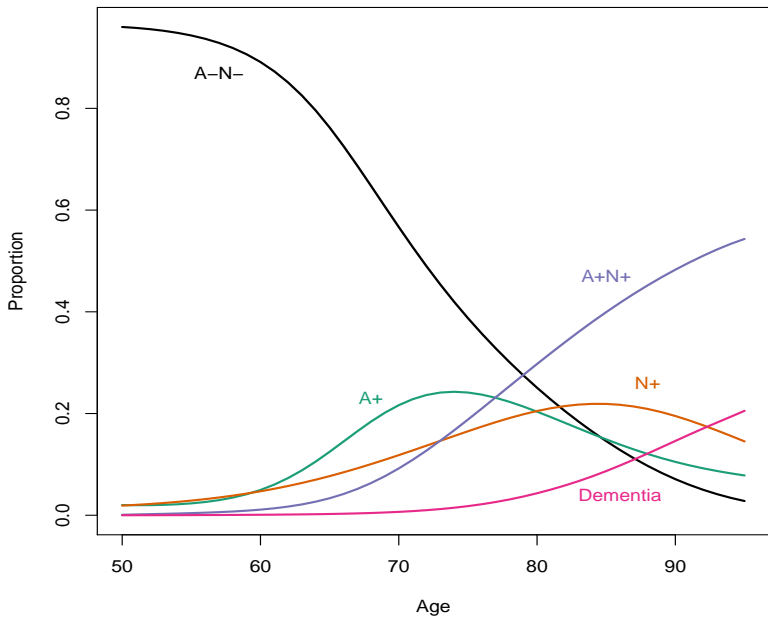












## Results

- ▶ Rates

- ▶ What is the pattern of rates?
  - ▶ The  $T0 \rightarrow T1$  rate is higher in the presence of A1, but not vice versa. (Amyloid deposits promote tau.)
  - ▶ A1/T1 promotes changes in N
- ▶ The role of covariates.
  - ▶ A positive APOE genotype affects A0/A1 transitions, but not others.
  - ▶ Other covariates affect N but not A or T

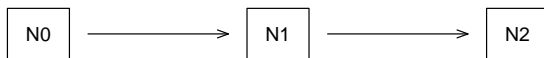
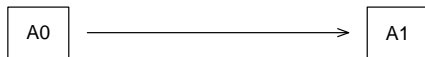
► Outcomes

- ▶ What is the probability of ever visiting the N2 state?
- ▶ What is the average duration of time spent in N2?
- ▶ What is the predicted fraction who go down each path?
- ▶ What is the impact of a change in one rate?

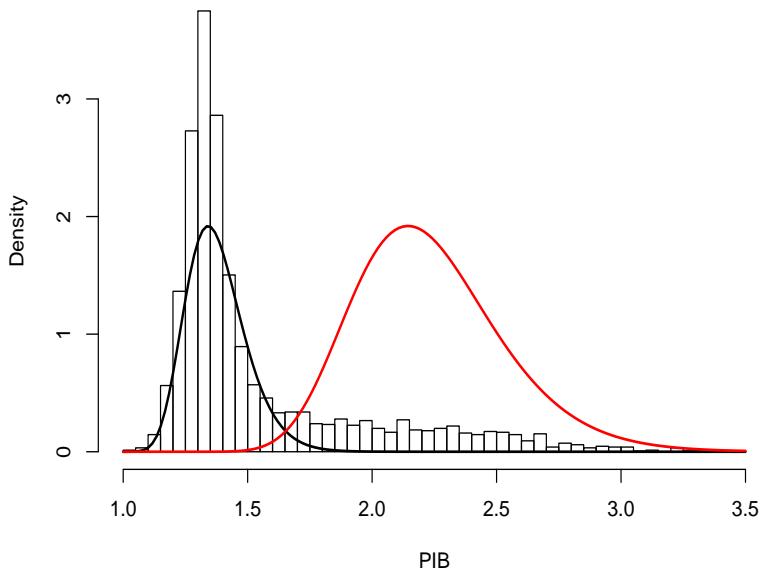
# Hidden Markov Model

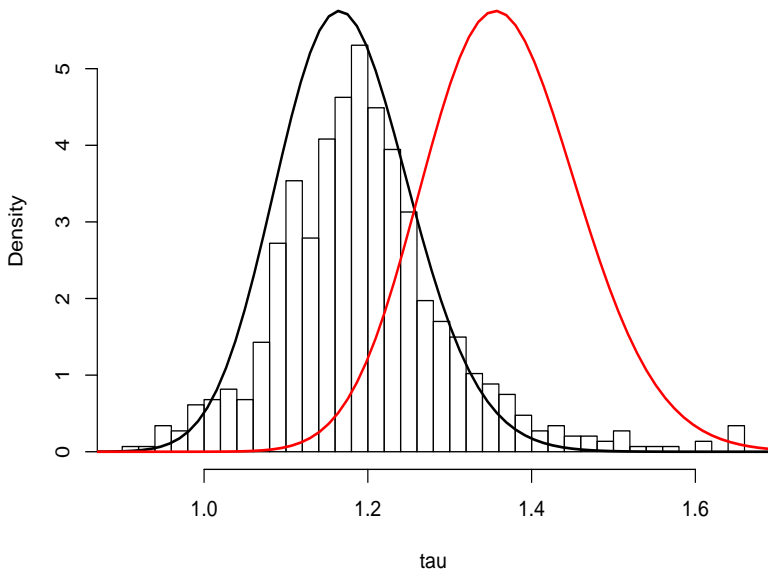
- ▶ The data consists of time, outcomes and covariates
  - ▶ The state is not observed directly, rather we see one or more outcomes that depend on the underlying state.
  - ▶ No need for (time1, time2, endpoint) notation
  - ▶ Data will have missing values, e.g., covariates on the day of death
- ▶ Same box and arrow model for the states, covariates connect to  $\lambda$  as before
- ▶ Another set of parameters for the arrows that connect state to outcome
- ▶ Allows for more episodic data.
- ▶ Much of the software is special purpose.

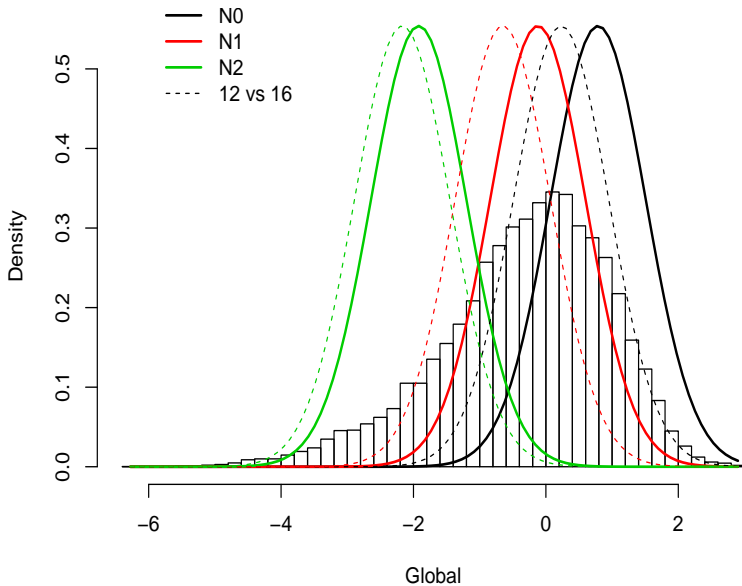




- ▶  $\log(\text{measured amyloid binding}) \sim N(A^-/A^+, \sigma)$
- ▶ global memory score  $\sim N(\mu, \tau)$   
 $\mu = \beta_0 + \beta_1 N + \beta_2 \text{sex} + \beta_3 \text{education}$
- ▶  $A^- : A^+$  rate depends on APOE status, but on gender
- ▶  $N$  transition rates depend on  $A$  but not vice-versa









# Conclusions

- ▶ Multi-state data ranges from the simple to the complex
- ▶ Good tools are available
- ▶ You need more than just a hazard ratio
- ▶ There is wide opportunity for new methods and software

# Model checks and time-dependent covariates









# Time-dependent covariates

The Cox model likelihood is set up like a lottery.

At each death time:

- ▶ Risk set = set of subjects *at risk* for death  
(if they had died, we would have recorded it)
- ▶ Score =  $r = \exp(X\beta)$  = “number of tickets”
- ▶  $L = r_d / \sum r_i = \text{Pr}(\text{the person who won, should have won})$
- ▶ Choose  $\beta$  to maximize  $L$

$$C = \frac{\sum_{\text{deaths}} \text{number with lower score } r \text{ than the winner}}{\sum_{\text{deaths}} \text{number at risk}}$$











## How to do it wrong



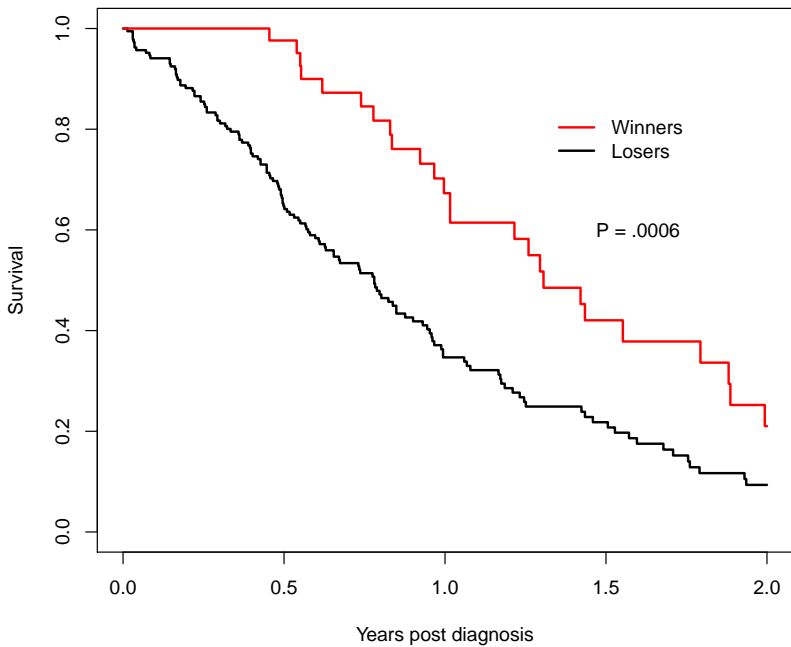






# Simulation

- ▶ Advanced lung cancer data set
- ▶ Assume bi-monthly visits
- ▶ Randomly mark 5% of the subjects “winners” at each visit.



The surprise is how *big* the error can be.

- ▶ Stanford Heart, time dependent transplant: .99 (.54, 1.8)
- ▶ Stanford Heart, ever transplant: .17 (.10, .28)
- ▶ Gail, "Does cardiac transplantation prolong life? A reassessment." Ann Int Med 1982.

# Cumulative dose received

- ▶ Does dose reduction harm patients?
- ▶  $x =$  cumulative dose received



# Cumulative dose received

- ▶ Does dose reduction harm patients?
- ▶  $x$  = cumulative dose received
- ▶  $x$  = fraction of expected

# Cumulative dose received

- ▶ Does dose reduction harm patients?
- ▶  $x$  = cumulative dose received
- ▶  $x$  = fraction of expected
- ▶  $x$  = fraction of expected, to date
- ▶ Redmond, Cancer Treatment Reports





# Prophetic variables

Some time-dependent covariates are not predictors of an event as much as they are markers of a failure-in-progress:

- ▶ Medication changes
  - ▶ Cessation of diuretics in heart failure
- ▶ Multiple-organ failure
- ▶ Ventilation
- ▶ “Have called the priest”
- ▶ Tautologies: lab test Tuesday, progression Wed
- ▶ Consider time delays

## Errors

- ▶ Placebo controlled trial
- ▶ Adverse reaction to active arm (rare, but expected)
- ▶ “Weeks on drug” add to the form



```

data new; set old;
  futime = fu_date - entry_dt;
  tstart =0;
  if (drug_weeks ne . and drug_weeks < futime/7) then do;
    * one of the crossover patients;
    tstop = drug_weeks *7; event = 0;    output;
    tstart= tstop; tstop=futime;
    arm =0; event = status; output;  *cross to placebo;
  end;
else do; * no crossover;
  tstop=futime; event=status; output;
end;

proc phreg (tstart, tstop) * status(0) = arm;

```

# Time delay

- ▶ Delayed labs
- ▶ UDCA
- ▶ Long term prediction
- ▶ Option in the `tmerge` function.





## Immortal time bias

Subjects are treated as 'at risk' when they actually are not.

- ▶ At risk: If the subject had had an event, we would have seen and recorded it.
- ▶ rhDNase study
- ▶ Time gaps in the REP
- ▶ Oscar winners live longer
- ▶ ever prescribed inhaled corticosteroids = asthmatic
- ▶ “at least 2 instances”, then use the first one

More

- ▶ Mark an adverse event as midway between visits
- ▶ Delete subjects who do not complete treatment
- ▶ Interpolate a lab value
- ▶ Add “average death rate” as a covariate
- ▶ Multi-state models create new opportunities for error
- ▶ ...

More

- ▶ Mark an adverse event as midway between visits
- ▶ Delete subjects who do not complete treatment
- ▶ Interpolate a lab value
- ▶ Add “average death rate” as a covariate
- ▶ Multi-state models create new opportunities for error
- ▶ ...
- ▶ Process
  - ▶ Think through special cases
  - ▶ Create the (start, stop] data set
  - ▶ Print out a portion and *read* it.
  - ▶ Think, pause, think
  - ▶ If results are too good to be true ...

# Summary

Time dependent covariates are a wonderful tool.

1. You must not look into the future.

# Summary

Time dependent covariates are a wonderful tool.

1. You must not look into the future.
2. Avoid prophetic variables.

# Summary

Time dependent covariates are a wonderful tool.

1. You must not look into the future.
2. Avoid prophetic variables.
3. It's all too easy to look into the future.

# Summary

Time dependent covariates are a wonderful tool.

1. You must not look into the future.
2. Avoid prophetic variables.
3. It's all too easy to look into the future.
4. Duration or rate variables work surprisingly rarely.



# Summary

Time dependent covariates are a wonderful tool.

1. You must not look into the future.
2. Avoid prophetic variables.
3. It's all too easy to look into the future.
4. Duration or rate variables work surprisingly rarely.
5. Bad things happen if you look into the future.

# Summary

Time dependent covariates are a wonderful tool.

1. You must not look into the future.
2. Avoid prophetic variables.
3. It's all too easy to look into the future.
4. Duration or rate variables work surprisingly rarely.
5. Bad things happen if you look into the future.
6. Short term prediction is uninteresting.

# Summary

Time dependent covariates are a wonderful tool.

1. You must not look into the future.
2. Avoid prophetic variables.
3. It's all too easy to look into the future.
4. Duration or rate variables work surprisingly rarely.
5. Bad things happen if you look into the future.
6. Short term prediction is uninteresting.
7. It is challenging to draw survival curves.

# Survival curves

- ▶ By definition, a survival curve is a look into tomorrow, given *what you know today*.
- ▶ In a time dependent model, looking ahead requires knowledge of future covariates.
  - ▶ you don't know them
  - ▶ so you have to guess
  - ▶ This is very, very hard to do effectively.

