Multi-state models and competing risks

Terry M Therneau
Mayo Clinic

July 1, 2015

1 Multi-state survival curves

Consider the three simple models in figure 1. Each box is a patient state and each arrow a possible transition. The top left figure is simple survival: all patients start in the alive state and can make a single transition to death. The top right depicts classic competing risks: all subjects start on the left, and each can make a single transition to one of 3 terminal states. The bottom figure shows a simple multi-state situation known as the illness-death model.

Traditionally the first case is handled by the Kaplan-Meier esimate and the second by the "cumulative incidence", the third case requires use of the the Aalen-Johansen estimate, which includes each of the first two as a special case. The AJ estimate is very flexible: subjects can appear in more than one state during the course of a study, subjects can start after time 0 (delayed entry), and they can start in any of the states. The `survfit` function implements the AJ estimate and can handle all these cases.

Let $A(t)$ be a matrix of cumulative hazard functions, whose ij element is the estimated cumulative hazard for transitions from state i to state j.

$$A_{ij}(t) = \int_0^t dN_{ij}(t)/r_i(t)$$

where dN counts the transitions and r is the number of subjects still at risk in a state. The diagonal elements of A are filled in last such that row sums of A are equal to zero. Then the Aalen-Johansen transition matrix is

$$P(t) = \prod_{s \le t} [I + dA(s)] \tag{1}$$

The product is over all time points $s \le t$ at which a transition occured, and dA is the change in the A matrix at that time point. For the two state model it is fairly easy to show that this reduces to the Kaplan-Meier.

The ith row of $P(t)$ estimates the fraction of subjects in each state at time t, given that subjects started in state i. The solution obeys the obvious constraint that the row sums at any time are equal to 1: each person has to be somewhere. I will refer to the resulting values as *prevalence* estimates. If there is no censoring then prevalence is particularly easy: at a given time just count the fraction of subjects in each state.

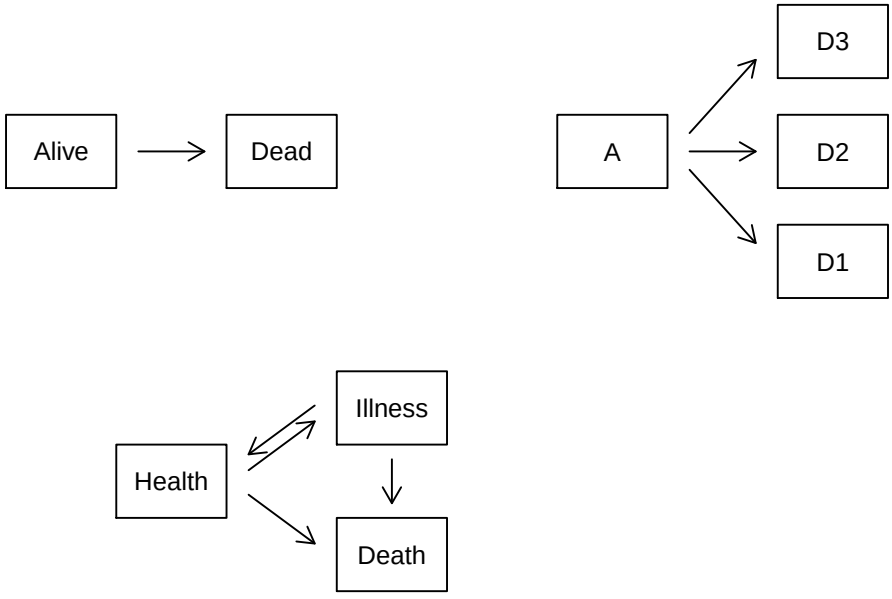


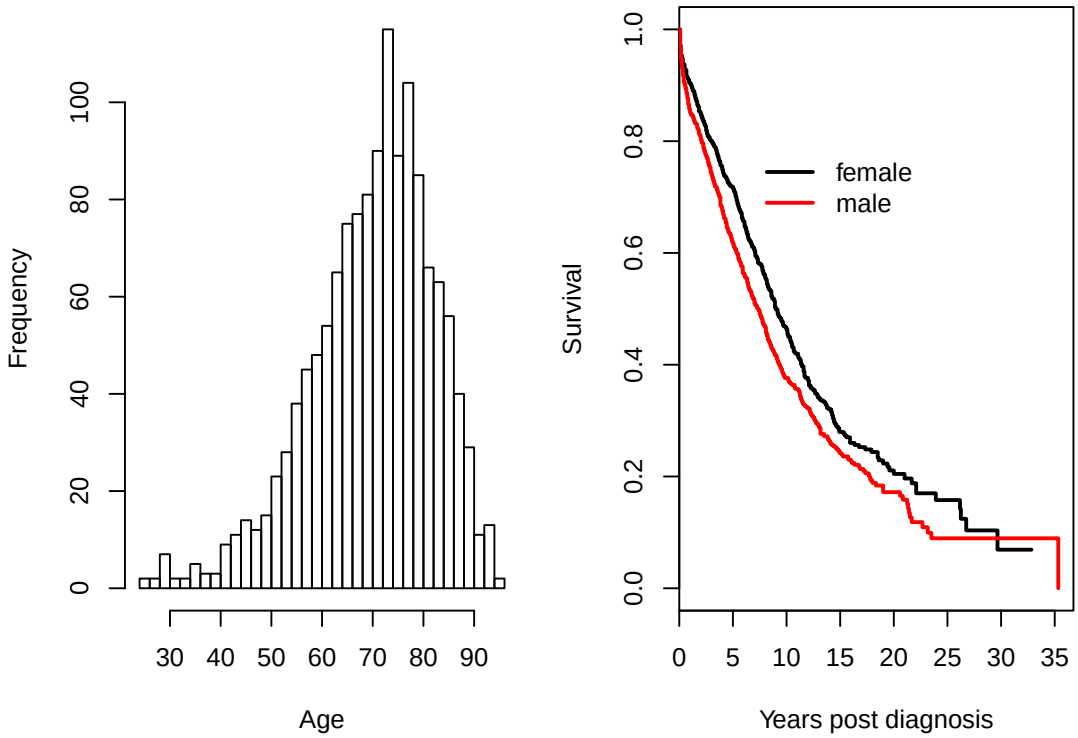
Figure 1: Three multi-state models. In the upper left is simple survival, in the upper right an example of competing risks, with the multi-state illness-death model below them.

In the simple two state model the prevalence in the alive state is the usual KM survival estimate, and we have P(alive) = 1 - P(dead). For simple survival we have gotten used to the idea of using P(dead) and 1- P(dead) interchangeably, but that habit needs to be left behind for multi-state models, for them the values $1 - P_k$ = probability(any other state than k) are not very useful. Plots for the 2 state case sometimes choose to show P(alive) and sometimes P(dead). Which one is used often depends on a historical whim of the disease specialty; cardiology journals for instance quite often use P(event) resulting in curves that rise starting from zero, but oncology journals invariably use P(alive) giving curves that fall downhill from 1. The survfit routine's historical default for the 2 state case is to print and plot P(alive), which reflects that the author of the routine was working primarily in cancer trials at the time said default was chosen. In the multi-state case, however, the curve for the initial state (leftmost in my diagrams) is rarely included in the final plot and curves start at 0.

Here is an example using a simple competing risks problem. The `mgus2` data set contains the time to plasma cell malignancy (PCM) and/or death for 1384 subjects diagnosed with monoclonal gammopathy of undetermined significance (MGUS). Survival and progression time are in months. The curve below shows ordinary Kaplan-Meier survival for these subjects, the mean age at diagnosis is just over 70 years.

```
> oldpar <- par(mfrow=c(1,2))
> hist(mgus2$age, nclass=30, main='', xlab="Age")
> with(mgus2, tapply(age, sex, mean))
       F        M
71.32171 69.67065
> mfit1 <- survfit(Surv(futime, death) ~ sex, data=mgus2)
> mfit1
Call: survfit(formula = Surv(futime, death) ~ sex, data = mgus2)

        n events median 0.95LCL 0.95UCL
sex=F 631    423    108     100     121
sex=M 753    540     88      79      97
> plot(mfit1, col=c(1,2), xscale=12, mark.time=FALSE, lwd=2,
       xlab="Years post diagnosis", ylab="Survival")
> legend(6, .8, c("female", "male"), col=1:2, lwd=2, bty='n')
> par(oldpar)
```

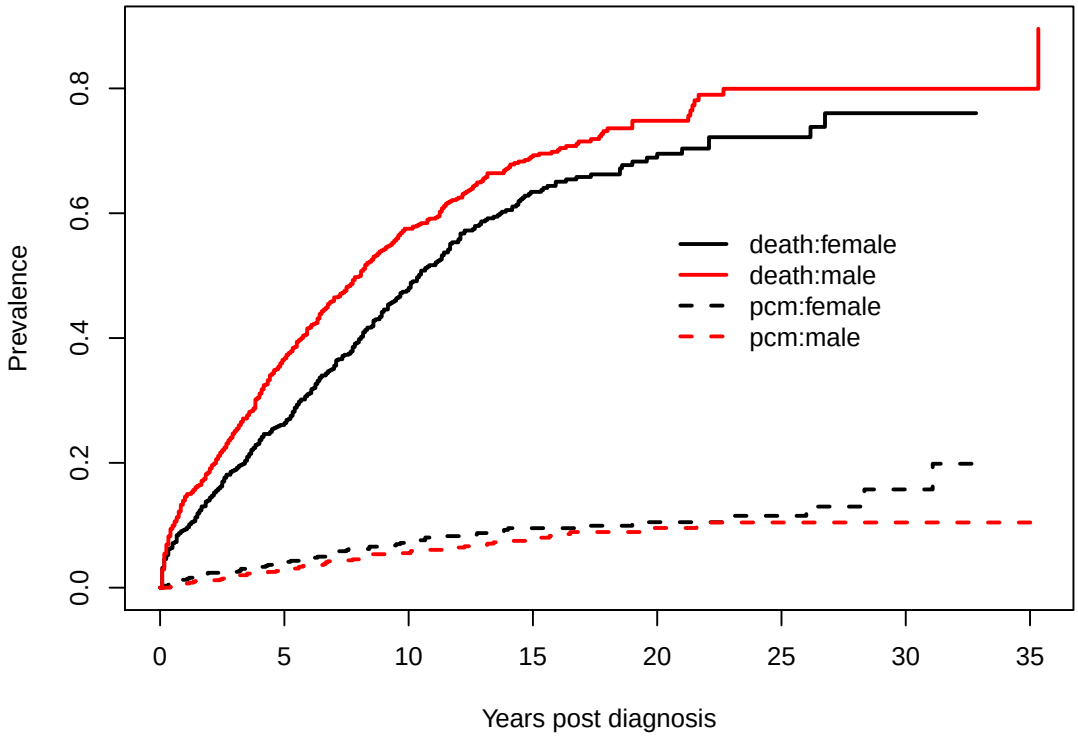


A second model for these subjects is competing risks, which corresponds to our second figure above. For this model we are only interested in the first event for each subject. Formally we are treating progression to a plasma cell malignancy (PCM) as an *absorbing state*, i.e., one that subjects never exit. We create a variable `etime` containing the time of the first progression, death, or last follow-up along with an event variable that contains the outcome. The starting data set `mgus2` has two pairs of variables `(ptime, pstat)` that contain the time to progression and `(futime, status)` that contain the time to death or last known alive, ignoring progression.

```
> etime <- with(mgus2, ifelse(pstat==0, futime, ptime))
> event <- with(mgus2, ifelse(pstat==0, 2*death, 1))
> event <- factor(event, 0:2, labels=c("censor", "pcm", "death"))
> table(event)
event
censor    pcm  death
   409    115    860
> mfit2 <- survfit(Surv(etime, event) ~ sex, data=mgus2)
> mfit2
Call: survfit(formula = Surv(etime, event) ~ sex, data = mgus2)

                n events median 0.95LCL 0.95UCL
pcm, sex=F    631    429     NA      NA      NA
death, sex=F  631    429    125     140     114
```

```
pcm, sex=M   753    546     NA      NA      NA
death, sex=M 753    546     97     106      84
> plot(mfit2, col=c(1,1,2,2), lty=c(2,1,2,1),
       xscale=12, mark.time=FALSE, lwd=2,
       xlab="Years post diagnosis", ylab="Prevalence")
> legend(20, .6, c("death:female", "death:male", "pcm:female", "pcm:male"),
         col=c(1,2,1,2), lty=c(1,1,2,2), lwd=2, bty='n')
```



The `mfit2` call is nearly identical to that for an ordinary Kaplan-Meier, with the exception of the `event` variable.

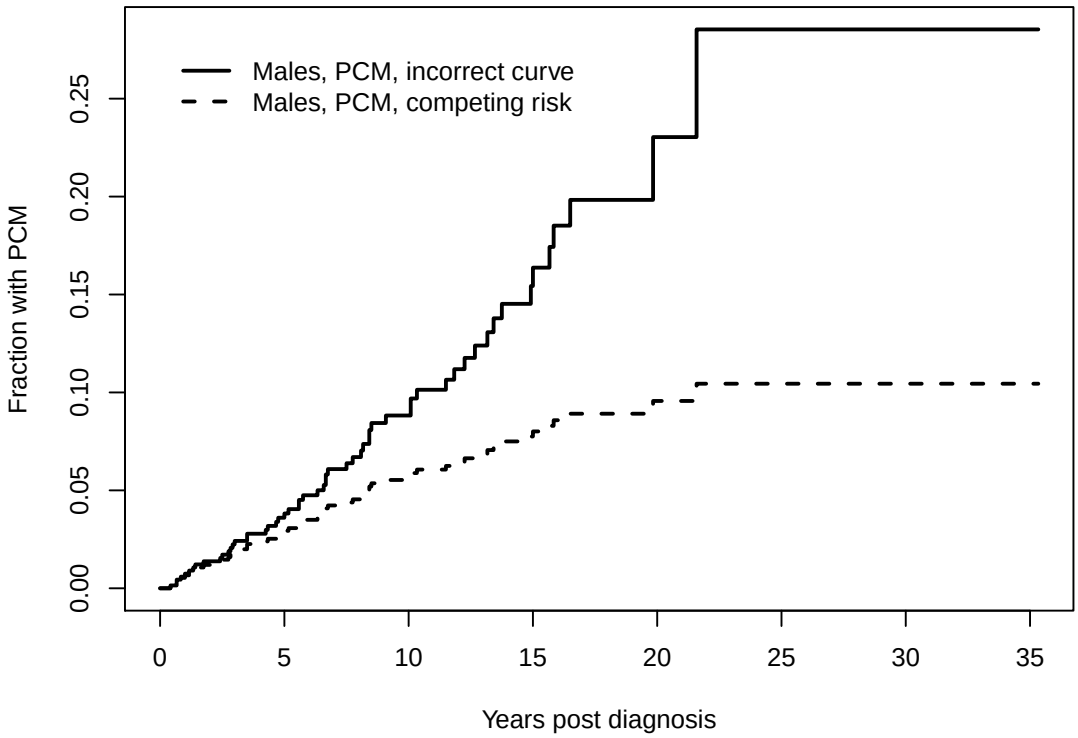
1. The event variable was created as a *factor*; whereas for ordinary single state survival the status is either 0/1 or TRUE/FALSE. The first level of the factor must be censoring, which is the status code for those whose follow-up terminated without reaching either endpoint. Codes for the remaining states can be in any order. The labels for the states are unrestricted.

2. A simple print of the `mfit1` 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.

3. Since these are prevalence estimates, the curves start at 0.

A common mistake with competing risks is to use the Kaplan-Meier separately on each event type while treating other event types as censored. The next plot is an example of this for the PCM endpoint.

```
> pcmbad <- survfit(Surv(etime, pstat) ~ sex, data=mgus2)
> plot(pcmbad[2], mark.time=FALSE, lwd=2, fun="event", conf=FALSE, xscale=12,
       xlab="Years post diagnosis", ylab="Fraction with PCM")
> lines(mfit2[2,1], lty=2, lwd=2, mark.time=FALSE, conf=FALSE, xscale=12)
> legend(0, .28, c("Males, PCM, incorrect curve", "Males, PCM, competing risk"),
        col=1, lwd=2, lty=c(1,2), bty='n')
```



There are two problems with the `pcmbad` fit. The first is that it attempts to estimate the expected rate of plasma cell malignancy if all other causes of death were disallowed. In this hypothetical world it is indeed true that many more subjects would progress to PCM, but it is not a world that any of us will ever inhabit and so is of questionable interest. The second problem is that the computation for this hypothetical case is only correct if all of the competing endpoints are independent, a situation which is almost never true. The competing risks curve estimates the fraction of MGUS subjects who will actually experience PCM, sometimes known as the lifetime risk.

The above code chose to plot only a subset of the curves, something that is often desirable in competing risks problems to avoid a "tangle of yarn" plot that simply has too many elements. This is done by subscripting the survfit object. For subscripting, multistate curves appear as a matrix with the outcomes as the second subscript. They are in order of the levels of `event`, i.e., as displayed by our earlier call to `table(event)`. The first subscript indexes the groups formed by the right hand side of the model formula, and will be in the same order as simple survival curves. Thus `mfit[2,1]` corresponds to males and the pcm endpoint.

A third example using the MGUS data treats it as a multi-state model. In this version a

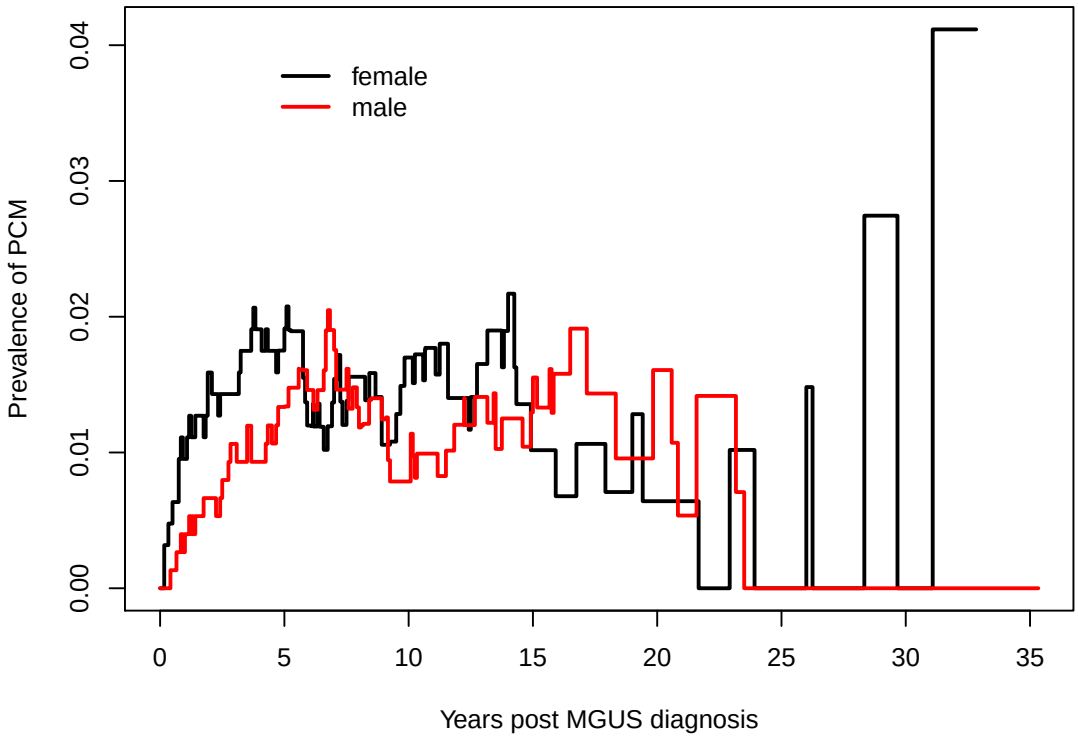
subject can have multiple transitions and thus multiple rows in the data set, and it is necessary to identify which data rows go with which subject via the `id` argument of `survfit` (valid estimates standard errors both depend on this). Our model looks like the illness-death model of figure 1 but with "plasma cell malignancy" as the upper state and no arrow for a return from that state to health. The necessary data set will have two rows for any subject who has further follow-up after a PCM and one row for all others. The data set is created below using the `tmerge` function, which is discussed in detail in another vignette.

We need to decide what to do with the 9 subjects who have PCM and death declared at the same time. They slipped through without comment in the earlier competing risks analysis, only when setting up this data set did I notice the ties. Looking back at the code, the prior example counted these subjects as a progression. In retrospect this is defensible: even though undetected before autopsy, the disease must have been present for some amount of time previous and so progression did occur first. For the multi-state model we need to be explicit in how this is coded since a sojourn time of 0 within a state is not allowed. Below we push the progression time back by .1 month when there is a tie, but that amount is entirely arbitrary.

```
> ptemp <- with(mgus2, ifelse(ptime==futime & pstat==1, ptime-.1, ptime))
> newdata <- tmerge(mgus2, mgus2,  id=id, death=event(futime, death))
> newdata <- tmerge(newdata, mgus2, id, pcm = event(ptemp, pstat))
> newdata <- tmerge(newdata, newdata, id, enum=cumtdc(tstart))
> with(newdata, table(death, pcm))
     pcm
death   0   1
    0 421 115
    1 963   0
```

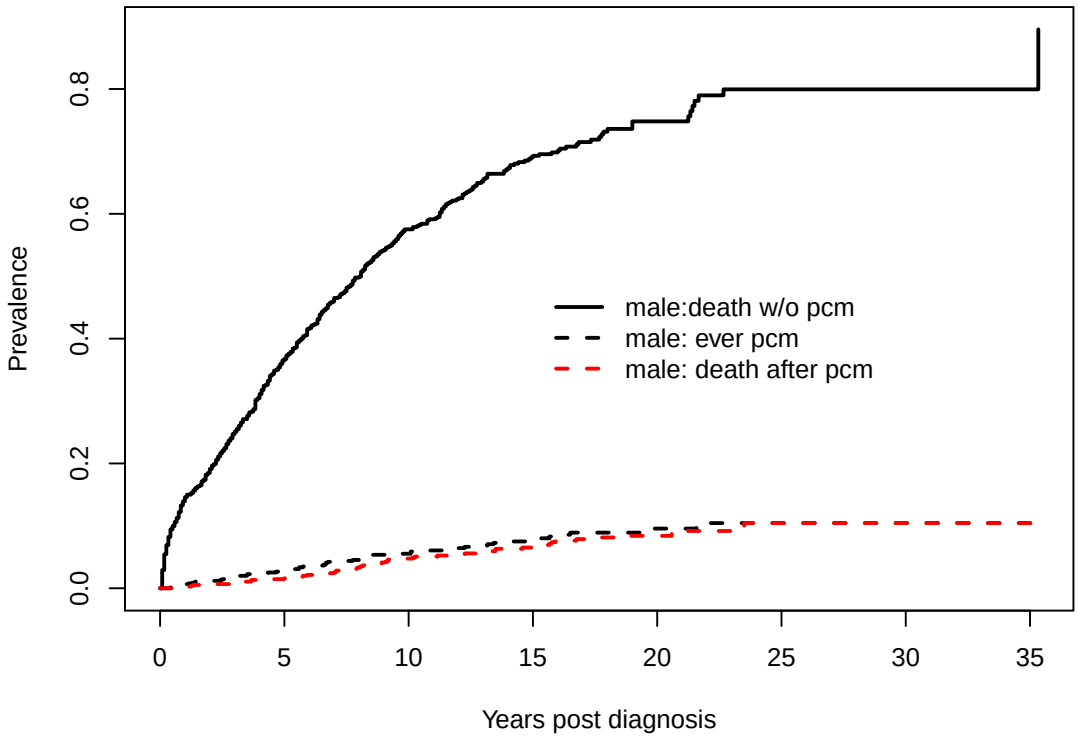
The table above shows that there are no observations in `newdata` that have both a pcm and death, i.e., the ties have been resolved. The last tmerge line above creates a variable `enum` which simply counts rows for each person; it will be used later.

```
> temp <- with(newdata, ifelse(death==1, 2, pcm))
> newdata$event <- factor(temp, 0:2, labels=c("censor", "pcm", "death"))
> mfit3 <- survfit(Surv(tstart, tstop, event) ~ sex, data=newdata, id=id)
> plot(mfit3[,1], mark.time=FALSE, col=1:2, lty=1, lwd=2,
       xscale=12,
       xlab="Years post MGUS diagnosis", ylab="Prevalence of PCM")
> legend(4, .04, c("female", "male"), lty=1, col=1:2, lwd=2, bty='n')
```



This plot is quite different in that it shows the fraction of subjects *currently* in the PCM state. Looking at the lower scenario in figure 1, this is the fraction of subjects in the upper right box. The curve goes up whenever someone enters the box and down when they leave. Myeloma survival was quite short during the era of this study and the proportion in the PCM state rarely rises above 2 percent. I have often found the three curve display below useful in these cases. It combines the results from competing risk model used above along with a second fit that treats death after PCM as a separate state from death before progression. Only males are shown in the plot to minimize overlap.

```
> d2 <- with(newdata, ifelse(enum==2, 4, as.numeric(event)))
> e2 <- factor(d2, labels=c("censor", "pcm", "death w/o pcm",
                            "death after pcm"))
> mfit4 <- survfit(Surv(tstart, tstop, e2) ~ sex, data=newdata, id=id)
> plot(mfit2[2,], lty=c(2,1),
       xscale=12, mark.time=FALSE, lwd=2,
       xlab="Years post diagnosis", ylab="Prevalence")
> lines(mfit4[2,3], mark.time=FALSE, xscale=12, col=2, lty=2, lwd=2,
        conf=FALSE)
> legend(15, .5, c("male:death w/o pcm", "male: ever pcm",
                   "male: death after pcm"), col=c(1,1,2), lty=c(1,2,2),
               lwd=2, bty='n')
```



When using multi-state data to create Aalen-Johansen estimates individuals are not allowed to have gaps in the middle of their time line. For example a data with (0, 30, pcm) and (50,70, death) as the two observations for a subject; the time from 30-70 is not accounted for. The method also does not account for what is known as panel data, where a subject's state is recorded at some a priori time such as a physician visit but the actual times of state transitions are unknown. Such data requires further assumptions about the transition process in order to model the outcomes, see for instance the msm package.

2 Models

For simple two-state survival the Cox model leads to three relationships

$$\lambda(t) = \lambda_0(t)e^{X\beta} \tag{2}$$

$$\Lambda(t) = \Lambda_0(t)e^{X\beta} \tag{3}$$

$$S(t) = \exp(-\Lambda(t)) \tag{4}$$

where λ, Λ and S are the hazard, cumulative hazard and survival functions, respectively. There is a single linear predictor which governs both the rate λ (the arrow in figure 1) and the prevalence value of the left hand box S. For multi-state models this simplicity no longer holds: proportional hazards does not lead to proportional prevalence curves.

2.1 Competing risks, Cox model

The Cox model approach starts by fitting separate models to each of the transitions. We will illustrate using the MGUS example.

```
> mtemp <- mgus2
> mtemp$age <- mtemp$age/10   #age in decades (easier coefficients)
> mtemp$etime <- etime
> mtemp$event <- event
> options(show.signif.stars = FALSE)  # display intelligence
> cfit2 <- coxph(Surv(futime, death) ~ age + sex + mspike, data=mtemp)
> cfit2
Call:
coxph(formula = Surv(futime, death) ~ age + sex + mspike, data = mtemp)


         coef exp(coef) se(coef)     z      p
age    0.6221    1.8629   0.0343 18.15 <2e-16
sexM   0.3536    1.4241   0.0659  5.37  8e-08
mspike 0.0311    1.0316   0.0598  0.52    0.6

Likelihood ratio test=393  on 3 df, p=0
n= 1373, number of events= 957
   (11 observations deleted due to missingness)
```

The effect of age and sex on non-PCM mortality is profound, which is not a surprise given the median starting age of 72. Risk rises 1.9 fold per decade of age and the death rate for males is 1.4 times as great as that for females. The size of the serum monoclonal spike is of no consequence for this endpoint either statistically or clinically.

```
> cfit1 <- coxph(Surv(ptime, pstat) ~ age + sex + mspike, mtemp)
> cfit1
Call:
coxph(formula = Surv(ptime, pstat) ~ age + sex + mspike, data = mtemp)


           coef exp(coef) se(coef)     z       p
age     0.16352   1.17765  0.08374  1.95   0.051
sexM   -0.00503   0.99498  0.18845 -0.03   0.979
mspike  0.88408   2.42075  0.16521  5.35 8.7e-08

Likelihood ratio test=30.3  on 3 df, p=1.22e-06
n= 1373, number of events= 115
   (11 observations deleted due to missingness)
> quantile(mgus2$mspike, na.rm=TRUE)
  0%  25%  50%  75% 100%
 0.0  0.6  1.2  1.5  3.0
```

The mspike size has a major impact on progression, however; each 1 gram change increases risk by 2.4 fold. The interquartile range of `mspike` is 0.9 gram so this risk increase is clinically important. The effect of age on the progression rate is much less pronounced, with a coefficient only 1/4 that for mortality, while the effect of sex on progression is negligible.

Notice that we did not do anything special to the data set or event codes for the Cox model. The focus of coxph is on the event rates, for which the correct denominator is the set of all subjects still at risk. This is exactly what is encoded by the (futime, death) and (ptime, pstat) pairs.

The effect of sex on the *lifetime* probability of PCM is not zero, however. Because of a longer lifetime, an average female with MGUS will spend more total years at risk for PCM than the average male, and so has a larger lifetime risk of PCM. The average rate of progression is about 1% per year, as shown below, while the average post diagnosis lifetime is 18 months longer for females.

```
> pfit1 <- pyears(Surv(ptime, pstat) ~ sex, mtemp, scale=12)
> round(100* pfit1$event/pfit1$pyears, 1)  # PCM rate per year
  F   M
1.1 1.0
> temp <- summary(mfit1, rmean="common")  #print the mean survival time
> round(temp$table[,1:6], 1)
      records n.max n.start events *rmean *se(rmean)
sex=F     631   631     631    423  142.4        6.1
sex=M     753   753     753    540  123.7        5.4
```

Prevalence estimates from the multi-state model involve the matrix $A(t;x)$ of cumulative hazard estimates. The i, j off diagonal element of $A(t;x)$ is the the cumulative hazard $\lambda_{ij}(t;x)$ for the $i \to j$ transition, obtained from the fitted Cox model for that transition. These predicted hazards are formed for a chosen set of covariates x, e.g. in the model above we could for instance choose predicted transitions for a 72 year old male with an mspike value of 1.1. Predicted curves from a Cox model are *always* with respect to a particular hypothetical subject. The notion of a baseline hazard, i.e. the hazard for a subject with all covariates equal to zero, is sometimes of mathematical convenience but only rarely corresponds to any patient of interest. The diagonal elements of A are filled in last and are chosen such that row sums of are 0.

The obvious analog to the univariate survival curve in equation (4) is the matrix exponential.

$$P(t;x) = e^{A(t;x)}$$

However, this computational approach is valid only if the A matrix is separable, i.e., $A(t;x) = A(x)g(t)$, something that holds true if there are no time dependent covariates in the model and if all the transitions share the same baseline hazard: a very unusual case. The matrix exponential formulation is fundamental to multi-state models with constant hazard however, see for instance the vignette for the `msm` package. For the Cox model we use the Aalen-Johansen estimator — the same approach used by `survfit` for non-parametric estimates.

$$P(t;x) = \prod_{s \le t} (I + dA(s;x)) \tag{5}$$

where the term dA is the increment in A at time s, and there is an increment at each event time. As with survival curves from an ordinary Cox model, any such curve is computed for a prespecified set of covariate values x which must be chosen by the user.

For illustration we will compute the probabilities of PCM from the model for males and females under 4 cases: age of 60 vs 80 and a serum mspike of 0.5 vs 1.5; these last are the approximately the quartiles of age and mspike. Each of surv1 and surv2 below will contain 8 curves, for the 8 combinations of sex, age and mspike.

```
> tdata <- expand.grid(mspike=c(.5, 1.5), age=c(6,8), sex=c("F", "M"))
> surv1 <- survfit(cfit1, newdata=tdata)  # time to progression curves
> surv2 <- survfit(cfit2, newdata=tdata)  # time to death curves
```

The individual survival curves are not actually of interest, since each is a Cox model analog of the 'pcmbad' curve we criticised earlier. Instead, the cumulative hazard portion of the results are used to build an Aalen-Johansen estimate. The A matrix is particularly easy in the competing risk case: all rows but the first will be 0, since only the $1 \to 2$ and $1 \to 3$ transitions are possible. Elements of the resulting 3 by 3 matrix $P(t)$ are the probability of going from state i to state j, since everyone starts in state 1 we are only interested in the first row of P. A computational nuisance is that the `surv1` and `surv2` curves do not necessarily jump at the same time. We use the summary function to select values on a common time scale. (The `summary.survfit` function was original written to provide printed values at specified times, but turns out to be an easy way to pluck off values.)

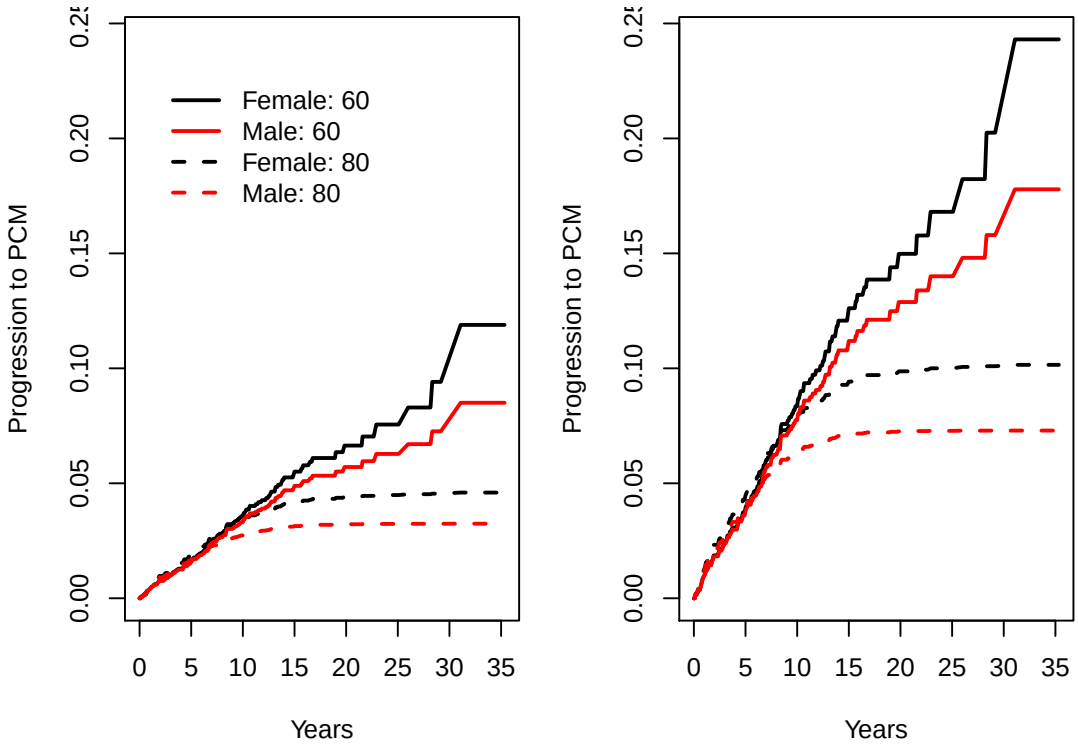
```
> cifun <- function(surv1, surv2) {
      utime <- sort(unique(surv1$time, surv2$time))
      jump1 <- diff(c(0, summary(surv1, times=utime, extend=TRUE)$cumhaz))
      jump2 <- diff(c(0, summary(surv2, times=utime, extend=TRUE)$cumhaz))
      dA  <- diag(3)
      prev  <- matrix(0., nrow= 1+length(utime), ncol=3)
      prev[1,1] <- 1  #initial prevalence at time 0: all are in the left box
      for (i in 1:length(utime)) {
          dA[1,2] <- jump1[i]  #fill in the first row of dA(s)
          dA[1,3] <- jump2[i]
          dA[1,1] <- 1- (jump1[i] + jump2[i])
          prev[i+1,] <- prev[i,] %*% dA
      }
      list(time=c(0, utime), P = prev)
  }
> # Get curves for the 8 cases, and save them in a matrix.
> #  Since they all come from the same pair of Cox models, the time values
> #  for all curves will be the same
> # The cifun function above is only designed to handle one of the 8 covariate
> #  patterns at a time, but survival curves can be subscripted.
> temp <- cifun(surv1[1], surv2[1])
> coxtime <- temp$time
> coxdeath <- coxpcm  <- matrix(0., nrow=length(temp$time), ncol=8)
> coxdeath[,1] <- temp$P[,3]
```

```
> coxpcm[,1]   <- temp$P[,2]
> for (i in 2:8){
      temp <- cifun(surv1[i], surv2[i])
      coxdeath[,i] <- temp$P[,3]
      coxpcm[,i]   <- temp$P[,2]
  }
> # Print out a M/F results at 20 years
> indx <- match(20*12, coxtime)
> progmat <- matrix(coxpcm[indx,], nrow=4)
> dimnames(progmat) <- list(c("a=50/ms=0.5", "a=50/ms=1.5",
                              "a=80/ms=0.5", "a=80/ms=1.5"),
                            c("female", "male"))
> round(100*t(progmat), 1)  #males and females at 20 years
       a=50/ms=0.5 a=50/ms=1.5 a=80/ms=0.5 a=80/ms=1.5
female         6.6        15.0         4.4         9.9
male           5.7        12.9         3.2         7.3
```

The above table shows that females are modeled to have a higher risk of 20 year progression, even though their hazard at any given moment is nearly identical to males. The difference at 20 years is on the order of our "back of the napkin" person-years estimate of 1% progression per year * 1.5 more years of life for the females, but the progression fraction varies substantially by group. Eighty year olds have a lower cumulative rate of PCM than 50 year olds due to a higher death rate, even though the hazard function for PCM rises with age.

A plot of the calculated progression curves is shown below. The left hand panel has predicted curves for those with a serum mspike of 0.5 and the right for mspike=1.5, and in all cases females are predicted to have a higher level of observed progression than males. Although the Cox model hazards are assumed to be proportional, the prevalence curves are not, however. For those diagnosed at an older age the prevalence curves flatten out after 10 years, simply because so few living subjects remain who are available to have a PCM event.

```
> par(mfrow=c(1,2))
> matplot(coxtime/12, coxpcm[,c(1,3,5,7)], col=c(1,1,2,2),
          lty=c(1,2,1,2), type='l', lwd=2, ylim=range(coxpcm),
          xlab="Years", ylab="Progression to PCM")
> legend(1, .23, c("Female: 60", "Male: 60", "Female: 80", "Male: 80"),
         lty=c(1,1,2,2), col=c(1,2,1,2), lwd=2, bty='n')
> matplot(coxtime/12, coxpcm[,c(2,4,6,8)], col=c(1,1,2,2),
          lty=c(1,2,1,2), type='l', lwd=2,
          xlab="Years", ylab="Progression to PCM")
```



In the competing risks case the prevalence function has an alternate form known as the *cumulative incidence* function

$$CI_k(t) = \int_0^t \lambda_k(u) S(u-) du \tag{6}$$

where λ_k is the incidence function for outcome k and S is the overall survival curve for "time to any endpoint". Proving that P_{1k} as computed by Aalen-Johansen estimate is equivalent to $CI(t)$ is straightforward. (The label "cumulative incidence" is one of the more unfortunate ones in the survival lexicon, since we normally use 'incidence' and 'hazard' as interchangeable synonyms but the CI is *not* a cumulative hazard.) For the general multi state case it is simplest to use the `mstate` package; it was designed for this task and will also compute appropriate confidence intervals. The latter are complex since they must account for the uncertainty in the underlying Cox models.

2.2 Fine-Gray model

For the competing risk case the Fine-Gray model provides an alternate way of looking at the data. As we saw above, the impact of a particular covariate on the final prevalence values P can be complex, even if the models for the hazards are relatively simple. Start with the functions $F_k(t) = P_{1k}(t)$, which can be thought of as the distribution function for the improper random variable $T^* = I(\text{endpoint} = k)T + I(\text{endpoint} \neq k)\infty$. Fine and Gray refer to F_k as a

subdistribution function. In an analog to the survival probability in the two state model define

$$\gamma_k(t) = -d \log[1 - F_k(t)]/dt I \tag{7}$$

and assume that $\gamma_k(t; x) = \gamma_{k0}(t) \exp(X\beta)$. In a 2 state model γ is the usual hazard function. In the same way that our multivariate Cox model `cfit2` made the simplifying assumption that the impact of male sex is to increase the hazard for death by a factor of 1.42 independent of the subject's age or serum mspike value, the Fine-Gray model assumes that each covariate's effect on $\log(1 - F)$ is a constant, independent of other variables. Both model's assumptions are wonderfully simplifying with respect to understanding a covariate — assuming of course that either assumption is correct. (In a multi-state model at least one of the two must be false.)

Let us look at the effect of sex on PCM using the Fine-Gray model, which can be computed using the `cmprsk` package. It does not use model formulas so variables need to be vectors or matrices.

```
> if (cmplib) {
      temp <- mtemp
      temp$fstat <- as.numeric(event)  # 1=censor, 2=pcm, 3=death
      temp$msex  <- with(temp, 1* (sex=='M'))
      fgfit1 <- with(temp, crr(etime, fstat, cov1= cbind(age, msex,  mspike),
                           failcode=2, cencode=1, variance=TRUE))
      fgfit2 <- with(temp, crr(etime, fstat, cov1=cbind(age, msex, mspike),
                            failcode=3, cencode=1, variance=TRUE))
      cmat <- rbind("FineGray: PCM" = fgfit1$coef,
                    "Cox: PCM" = coef(cfit1),
                    "FineGray: death" = fgfit2$coef,
                    "Cox: death" = coef(cfit2))
      round(cmat,2)
  }
11 cases omitted due to missing values
11 cases omitted due to missing values
                   age  msex mspike
FineGray: PCM   -0.17 -0.21   0.89
Cox: PCM         0.16 -0.01   0.88
FineGray: death  0.59  0.37  -0.15
Cox: death       0.62  0.35   0.03
```

The program has determined that female sex increases the PCM outcome by exp(.21) = 1.24 fold, for all values of age and mspike. The Cox model shows no effect of sex on the instantaneous hazard, but as shown in the last section Cox models do predict higher female prevalence. We had also seen that older subjects are less likely to experience PCM due to the competing risk of death; this is reflected in the FG model as a negative coefficient for age.

The primary strength of the Fine-Gray model with respect to the Cox model approach is that if "lifetime risk" is a primary question then the model has given us a simple and digestible answer to that question: females have a 1.2 fold higher risk. A primary problem of the model is that we can't go backwards: there is not a simple analog to the Aalen-Johansen estimator to carry one from F back to Λ. If one fits a set of Cox models to the arrows (hazards) then the

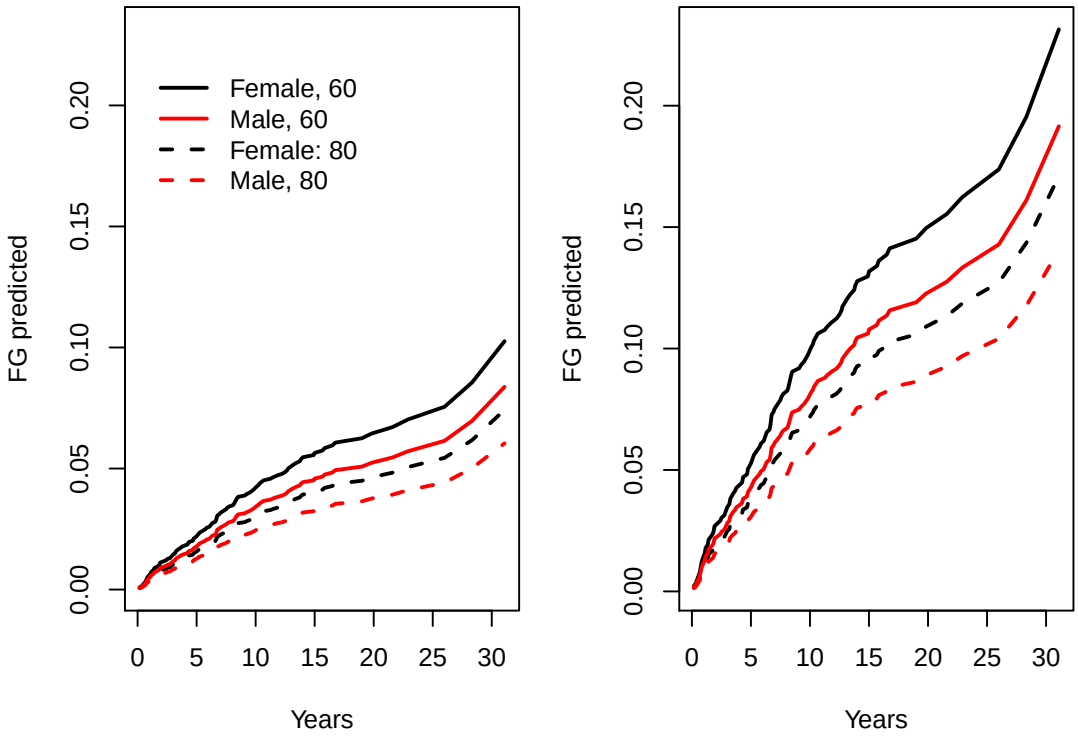
boxes (prevalence) of figure 1 can be examined post fit. With the Fine-Gray approach we have information only on the boxes.

To compare the two fits we can look at what the female/male ratios for each of our four chosen age/mspike combinations, when P is computed from the Cox models.

```
> cox.f <- log(1- progmat)    #log(1-P)
> round(cox.f[,1] / cox.f[,2], 2)
a=50/ms=0.5 a=50/ms=1.5 a=80/ms=0.5 a=80/ms=1.5
       1.17        1.18        1.38        1.38
```

The Cox models, which assume proportional hazards, show a larger subdistribution hazard for those who are older, those with higher mspike values, and at longer follow-up times. The overall average, however, is similar to the single value that results from a Fine-Gray model. The predicted curves are however nuch different from those shown before for a Cox model; the Fine-Gray curves are displayed below with predictions for mspike=0.5 on the left and 1.5 on the right.

```
> if (cmplib) {
  par(mfrow=c(1,2))
  fdata <- model.matrix(~age + sex + mspike, data=tdata)[,-1] #remove intercept
  fpred <- predict(fgfit1, cov1=fdata)
  matplot(fpred[,1]/12, fpred[,c(2,4,6,8)], col=c(1,1,2,2), lty=c(1,2,1,2),
          ylim=range(fpred[,-1]),
         type='l', lwd=2, xlab="Years", ylab="FG predicted")
  legend(0, .22, c("Female, 60", "Male, 60","Female: 80", "Male, 80"),
         col=c(1,2,1,2), lty=c(1,1,2,2), lwd=2, bty='n')
  matplot(fpred[,1]/12, fpred[,c(3,5,7,9)], col=c(1,1,2,2), lty=c(1,2,1,2),
         type='l', lwd=2, xlab="Years", ylab="FG predicted")
  }
```



This tells a very different story than the Cox model prevalence curves. Which is correct? Individual non-parametric prevalence curves are not as helpful as one would hope: there are simply too few progression events when separated into 8 groups.

```
> if (cmplib)
  fgfit3 <- with(temp, crr(etime, fstat, cov1= cbind(age, msex,  mspike),
                           failcode=2, cencode=1, variance=TRUE,
                           cov2=msex, tf = function(x) log(x)))
11 cases omitted due to missing values
```

A deeper analysis is called for, but will have to be left for another day.

3 Conclusions

When working with acute disease such as advanced cancer or end-stage liver disease there is often a single dominating endpoint. Ordinary single event Kaplan-Meier curves and Cox models are then efficient and sufficient tools for much of the analysis. Such data was the primary use case for survival analysis earlier in the author's career. Data with multiple important endpoints is now common, and multi-state methods are an important addition to the statistical toolbox. As shown above, they are now readily available and easy to use.

It is sometimes assumed that the presence of competing risks *requires* the use of a Fine-Gray model (I have seen it in referee reports), but this is not correct. The model may often be useful,

but is one available option among many. Grasping the big picture for a multi-state data set is always a challenge and we should make use of as many tools as possible. We are often minded of the story of a centerian on his 100th birthday proclaiming that he was looking forward to many more years because "I read the obituaries every day, and you almost never see someone over 100 there". It is not always easy to reason correctly from cumulative deaths back to hazard rates.

An advantage of the Cox model is that it has better diagnostic tools available, e.g., evaluation of the proportional hazards assumption via `cox.zph` or the martingale residuals, which can help to further refine our understanding. It is also easier to link hazard rates to a biologic rationale (perhaps incorrectly) which can help in explaining a data set.