Survival Analysis of Leukemia Treatment

Greg Johnson

8/14/2017

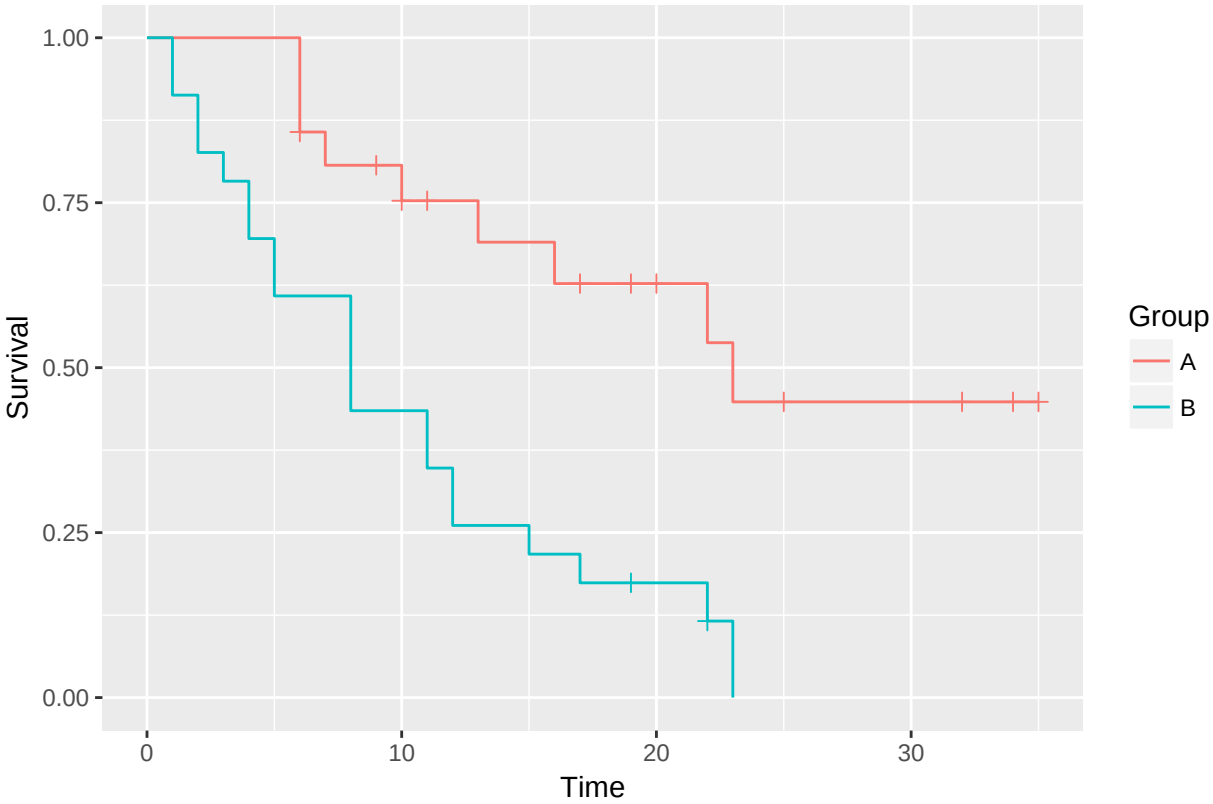
We have survival data on two types of treatment for Leukemia. We want to compare survival. In other words, we want to assess whether there is a benefit to being placed in the treatment group as opposed to the control group. There is no explicit variable for censoring but we can treat the indicator variable of death as a censoring variable (alive means censored). This will allow us to perform a proportional hazards model which allows us to compare hazard functions between people in the treatment group versus the control group after we have accounted for the effect of white blood cell count at initial diagnosis.Recall that the hazard function at time t is the probability of you dying *now* given that you've survived until this time.

```
Leuk<-read.csv("data/Leukemia.csv")
```

Before we blindly apply a model, let's look at the survival curves of the treatment group versus the control group. The curves will be unadjusted for logWBC, but the plot should give us some intuition about what our model may find.

```
ggsurv(survfit(Surv(Time,Result=="Death")~Group,data=Leuk))
```

```
## Loading required package: scales
```



It looks like there is a higher probability of living longer if you're in the treatment group versus the control group. Now let's fit the model.

```
cox.Leuk<-coxph(Surv(Time,Result=="Death")~Group + logWBC,data=Leuk)
```

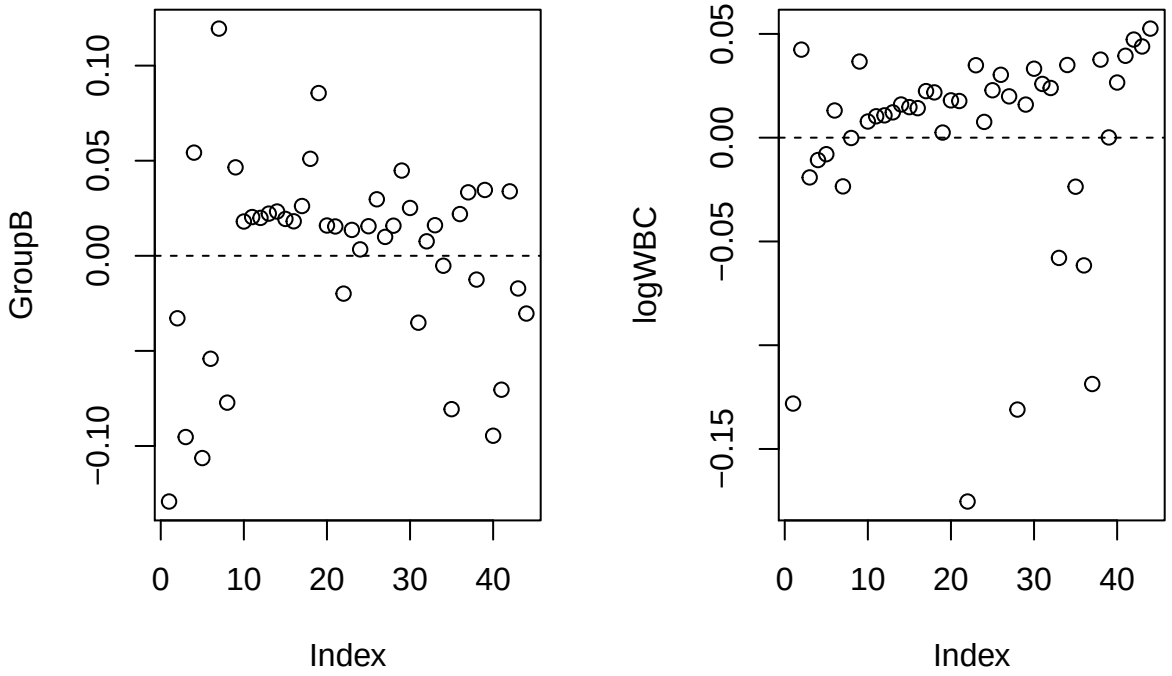
Let's look at some model diagnostics.

```
cox.zph(cox.Leuk)
```

```
##             rho   chisq     p
## GroupB -0.00804 0.00173 0.967
## logWBC  0.06443 0.17891 0.672
## GLOBAL       NA 0.18953 0.910
```

Our model assumption of proportional hazards is verified.

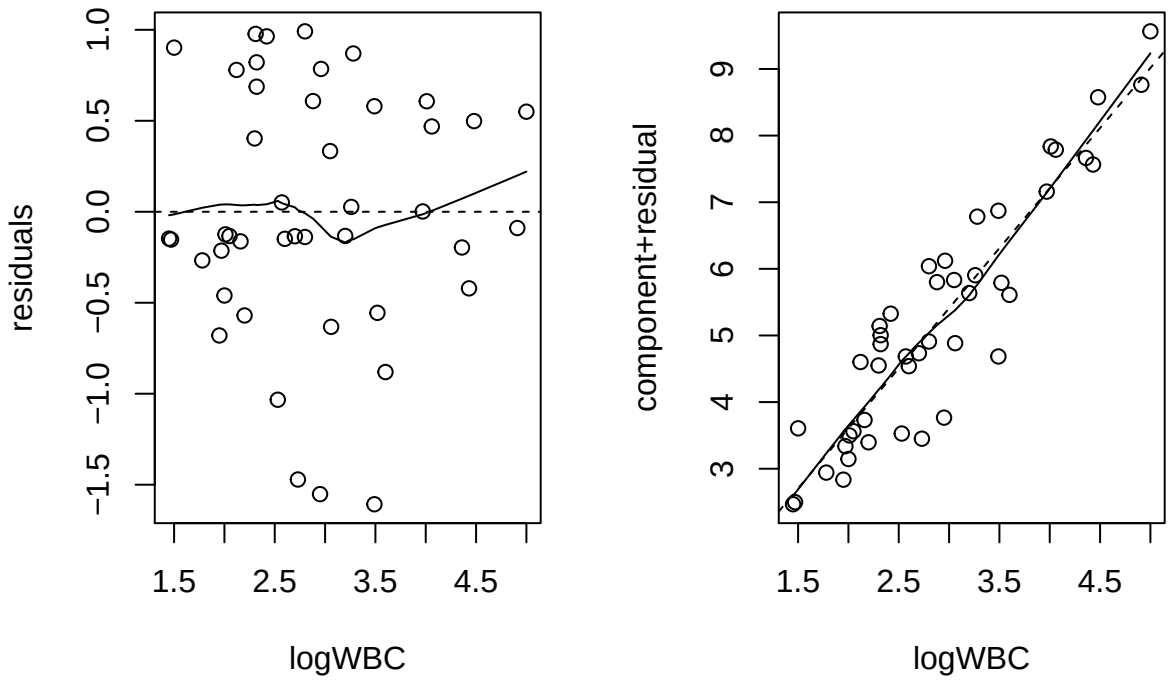
```
par(mfrow=c(1,2))
dfbeta<-residuals(cox.Leuk,type="dfbeta")
for(j in 1:2){
  plot(dfbeta[,j],ylab=names(coef(cox.Leuk))[j])
  abline(h=0,lty=2)
}
```



We have some cases that have a greater effect on the parameter estimate of the treatment effect and the parameter estimate of the wbc effect. However the magnitude of the effect is trivial when compared to the magnitude of the parameter estimates themselves. Therefore the data doesn't have any overly influential observations.

```
par(mfrow=c(1,2))
res<-residuals(cox.Leuk,type="martingale")
X<-as.matrix(Leuk[,"logWBC"])
plot(X,res,xlab="logWBC",ylab="residuals")
abline(h=0,lty=2)
lines(lowess(X,res,iter=0))

b<-coef(cox.Leuk)[2]
plot(X,b*X+res,xlab="logWBC",ylab="component+residual")
abline(lm(b*X+res~X),lty=2)
lines(lowess(X,b*X+res,iter=0))
```



Looks like log of white blood cell count meets our assumption of linearity. We can now safely interpret our model.

```
summary(cox.Leuk)
```

```
## Call:
## coxph(formula = Surv(Time, Result == "Death") ~ Group + logWBC,
##     data = Leuk)
##
##   n= 44, number of events= 30
##
##          coef exp(coef) se(coef)     z Pr(>|z|)
## GroupB 1.3281    3.7738   0.4285 3.099  0.00194 **
## logWBC 1.8029    6.0674   0.3298 5.467 4.59e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##        exp(coef) exp(-coef) lower .95 upper .95
## GroupB     3.774     0.2650     1.629     8.741
## logWBC     6.067     0.1648     3.179    11.581
##
## Concordance= 0.855  (se = 0.062 )
## Rsquare= 0.669   (max possible= 0.987 )
## Likelihood ratio test= 48.66  on 2 df,   p=2.711e-11
## Wald test            = 35.37  on 2 df,   p=2.087e-08
## Score (logrank) test = 48.93  on 2 df,   p=2.366e-11
```

If we set our alpha level at .05 - recall that this is essentially our comfort level for declaring that there is an effect of treatment - we can see that once we account for the effects of log WBC, there is a difference in hazard between the treatment group and the control group.

Exponentiating our parameter for the treatment effect, our model tells us that being in the control group is associated with a 3.774-factor increase in the hazard of death. This means that being in the control group, on average, increases hazard of death by 277%! On the flip side, this is a 73.5% decrease in hazard of death,

on average, if you are in the treatment.

Our model has found a statistically significant effect of treatment on hazard of death in our data and the estimated effect size is large and positive. If future replication studies verify this result, this treatment will have huge benefits for Leukemia patients.