

PSTAT175 – Project

Kartik Adimulam

June 2, 2025

1. Introduction

Overview of the Dataset

We are studying the Chemotherapy for Stage B/C colon dataset from the `survival` package in R (Laurie et al., 1989). This dataset comes from one of the first successful trials of chemotherapy for colon cancer. It examines the impact of treating patients with just a compound called Levamisole or a combination of it and 5-FU, a moderately toxic chemotherapy agent, or neither. Each patient has two records, one for recurrence and one for death, which we will take into account in our analysis.

Our dataset contains survival time data of days until death of colon cancer patients that received different levels of chemotherapy. The primary covariate in our analysis is the `rx` treatment variable which has 3 levels (no treatment, low-toxicity treatment, and moderate toxicity treatment) and the censoring status indicator.

There are other covariates such as:

- `sex`: male (1) or female (0)
- `age`: in years
- `obstruct`: was the colon obstructed by the tumor?
- `perfor`: was the colon perforated?
- `adhere`: is the tumor on nearby organs?
- `nodes`: how many lymph nodes have cancer?
- `differ`: is the tumor differentiated?
- `extent`: how much has the tumor spread locally?
- `surg`: how long from surgery to registration?
- `node4`: indicator variable of more than 4 positive lymph nodes
- `etype`: did subject's cancer recur or death

In our analysis we will use a significance level of $\alpha = 0.05$.

Motivating Questions

The scientific questions we are trying to answer are:

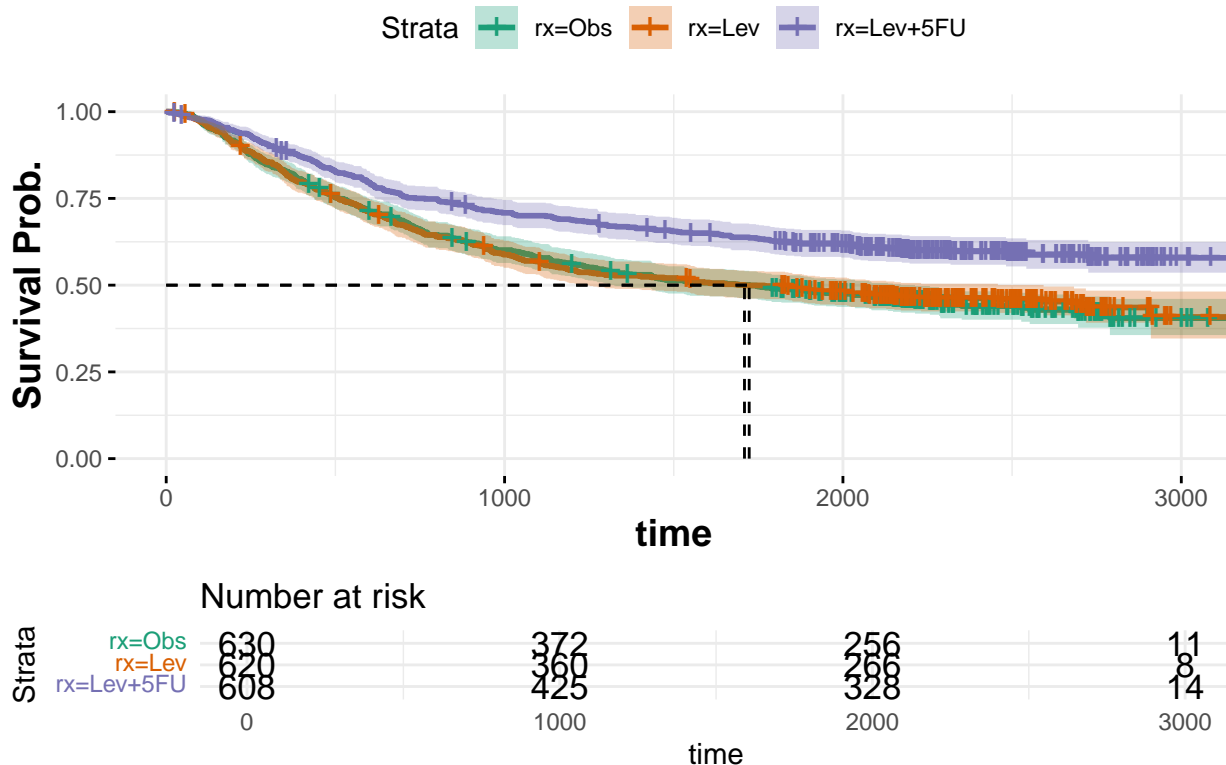
1. Does the type of treatment (levamisole or 5-FU) affect survival times?
2. Are there some covariates that are not strong predictors and if so which ones?
3. How much is survival probability affected by different covariate values?
4. Do covariates affect survival probability differently over time?
5. Are there important interactions between these covariates?

Survival Probability of Treatments

In our first step of evaluating which treatments affect survival times, we create a Kaplan-Meier estimate plot to visualize and gain a base understanding of how the 3 treatments in the dataset affect the survival probabilities of colon cancer patients.

```
colon=colon
colon_fit = survfit(Surv(time, status) ~ rx, colon)
ggsurvplot(colon_fit, colon, title='KM Estimator Colon Cancer',xlab='time',
  ylab='Survival Prob.',surv.median.line = 'hv',palette = "Dark2",
  conf.int = TRUE,
  risk.table = TRUE,
  ggtheme = theme_minimal(),
  font.main = c(16, "bold", "black"),
  font.x = c(14, "bold"),
  font.y = c(14, "bold"))
```

KM Estimator Colon Cancer



From the graph, we observe that the patients with the treatment that combines 5FU and Levamisole have a higher survival probability than those with no treatment and those with just the Levamisole treatment. The survival probability for patients with the Levamisole treatment seem to not have a large difference compared to those with no treatment, but the patients with no treatment have the lowest survival time.

Splitting Data

```
colon.recur = colon[colon$etype==1,]
colon.death = colon[colon$etype==2,]

colon.recur.cox = coxph(Surv(time,status)~rx, colon.recur)
summary(colon.recur.cox)

## Call:
## coxph(formula = Surv(time, status) ~ rx, data = colon.recur)
##
```

```
## n= 929, number of events= 468
##
##          coef exp(coef) se(coef)      z Pr(>|z|)
## rxLev      -0.01512   0.98499  0.10708 -0.141   0.888
## rxLev+5FU  -0.51209   0.59924  0.11863 -4.317 1.58e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          exp(coef) exp(-coef) lower .95 upper .95
## rxLev          0.9850      1.015   0.7985   1.2150
## rxLev+5FU      0.5992      1.669   0.4749   0.7561
##
## Concordance= 0.554 (se = 0.013 )
## Likelihood ratio test= 24.34 on 2 df,  p=5e-06
## Wald test              = 22.58 on 2 df,  p=1e-05
## Score (logrank) test = 23.07 on 2 df,  p=1e-05
```

```
anova(colon.recur.cox)
```

```
## Analysis of Deviance Table
## Cox model: response is Surv(time, status)
## Terms added sequentially (first to last)
##
##      loglik  Chisq Df Pr(>|Chi|)
## NULL -3040.3
## rx   -3028.1 24.343  2 5.175e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
colon.death.cox = coxph(Surv(time, status) ~ rx, colon.death)
summary(colon.death.cox)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ rx, data = colon.death)
##
## n= 929, number of events= 452
##
##          coef exp(coef) se(coef)      z Pr(>|z|)
## rxLev      -0.02664   0.97371  0.11030 -0.241  0.80917
## rxLev+5FU  -0.37171   0.68955  0.11875 -3.130  0.00175 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          exp(coef) exp(-coef) lower .95 upper .95
## rxLev          0.9737      1.027   0.7844   1.2087
## rxLev+5FU      0.6896      1.450   0.5464   0.8703
##
## Concordance= 0.536 (se = 0.013 )
## Likelihood ratio test= 12.15 on 2 df,  p=0.002
## Wald test              = 11.56 on 2 df,  p=0.003
## Score (logrank) test = 11.68 on 2 df,  p=0.003
```

```
anova(colon.death.cox)
```

```
## Analysis of Deviance Table
## Cox model: response is Surv(time, status)
```

```
## Terms added sequentially (first to last)
##
##      loglik  Chisq Df Pr(>|Chi|)
## NULL -2930.2
## rx    -2924.1 12.148  2    0.002302 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This analysis has two parts. Each person in the dataset has two entries. The first is when the person's cancer returns and the next is when the person passes away. We created two models, one for time until recurrence and one for time until death. Tests were then performed on the significance of treatment type compared to each other.

In our time until recurrence model, treatment using Levamisole compared with the baseline effect has a p -value of 0.888. This teaches us that this low-toxicity compound is not effective in reducing the time until recurrence for colon cancer patients.

However, the p -value of the treatment which combines 5FU (moderately toxic) treatment and Levamisole (low toxicity) is 0.00002. We can conclude that, at the given significance level, the combination of 5FU and Levamisole is effective in reducing time until recurrence but the treatment of just using Levamisole is not.

Overall, we observe using a likelihood ratio test p -value of ~ 0 that the treatment itself is effective in reducing time until recurrence, but most of this effect is likely due to the high effectiveness of Levamisole+5FU treatment.

In our time until death model, treatment using Levamisole compared with the baseline effect has a p -value of 0.809. This teaches us that this low-toxicity compound is not effective in reducing the time until death for colon cancer patients.

However, the p -value of the treatment which combines 5FU (moderately toxic) treatment and Levamisole (low toxicity) is 0.0018. We can conclude that the combination of 5FU and Levamisole is effective in reducing time until death but the treatment of just using Levamisole is not. This is an outcome that is discussed in the original study, which states that “after correction for imbalances in prognostic variables, were only suggestive for levamisole alone ($P = .05$) but quite significant for levamisole plus 5-FU ($P = .003$)” (Laurie et al., 1989). It is interesting, however, that our p -value for levamisole treatment alone obtained from our cox PH model is so much higher than the p -value obtained by the researchers of this study. Although both of us agree that Levamisole treatment in itself is ineffective, our level of ineffectiveness is much higher than that of the study. This could potentially be attributed to the treatment of other covariates by the researchers differing from our own treatment.

Overall, we observe using a likelihood ratio test p -value of 0.0023 that the treatment itself is effective in reducing time until death, but most of this effect is likely due to the high effectiveness of Levamisole+5FU treatment.

Confidence Intervals

```
colon.recur.CI = exp(confint(colon.recur.cox))
colon.death.CI = exp(confint(colon.death.cox))
colon.recur.CI
```

```
##              2.5 %    97.5 %
## rxLev        0.7985270 1.2149950
## rxLev+5FU    0.4749267 0.7560926
```

```
colon.death.CI
```

```
##              2.5 %    97.5 %
## rxLev        0.7844054 1.2087109
```

```
## rxLev+5FU 0.5463673 0.8702657
```

The confidence intervals for recurrence and death are shown above.

Again the CIs show that Levamisole treatment is not significant but Levamisole+5FU treatment is effective in reducing time until recurrence and death. Our hazard ratio CI for Lev treatment contains 1 for both models, but our hazard ratio CI for Lev+5FU treatment does not contain 1 for both models. We are 95% confident that Lev+5FU treatment significantly affects both time until recurrence and time until death.

Hazard Ratio

```
exp(coef(colon.recur.cox))
```

```
##      rxLev rxLev+5FU  
## 0.9849905 0.5992400
```

```
exp(coef(colon.death.cox))
```

```
##      rxLev rxLev+5FU  
## 0.9737142 0.6895540
```

At any point in time Lev+5FU treated patients are 0.6 times as likely to have a colon cancer recurrence compared to patients without treatment. The patients with Lev+5FU are 0.69 times as likely to have a colon cancer death compared to patients without treatment.

2. Model Fitting

Recurrence of Cancer Subset

Using forward stepwise selection and AIC, we can find the most relevant covariates to build and find the best model at predicting survival times. We first perform this process on the subset of data where there was a recurrence of cancer in patients.

```
colon.rx.sex = coxph(Surv(time,status)~rx + sex, colon.recur)
colon.rx.age = coxph(Surv(time,status)~rx + age, colon.recur)
colon.rx.obstruct = coxph(Surv(time,status)~rx + obstruct, colon.recur)
colon.rx.perfor = coxph(Surv(time,status)~rx + perfor, colon.recur)
colon.rx.adhere = coxph(Surv(time,status)~rx + adhere, colon.recur)
colon.rx.nodes = coxph(Surv(time,status)~rx + nodes, colon.recur)
colon.rx.differ = coxph(Surv(time,status)~rx + differ, colon.recur)
colon.rx.extent = coxph(Surv(time,status)~rx + extent, colon.recur)
colon.rx.surg = coxph(Surv(time,status)~rx + surg, colon.recur)
colon.rx.node4 = coxph(Surv(time,status)~rx + node4, colon.recur)
```

```
AIC(colon.rx.sex, colon.rx.age, colon.rx.obstruct, colon.rx.perfor,
    colon.rx.adhere, colon.rx.nodes, colon.rx.differ, colon.rx.extent,
    colon.rx.surg, colon.rx.node4)
```

##		df	AIC
##	colon.rx.sex	3	6060.887
##	colon.rx.age	3	6059.623
##	colon.rx.obstruct	3	6058.401
##	colon.rx.perfor	3	6060.226
##	colon.rx.adhere	3	6056.852
##	colon.rx.nodes	3	5827.396
##	colon.rx.differ	3	5898.618
##	colon.rx.extent	3	6030.710
##	colon.rx.surg	3	6056.936
##	colon.rx.node4	3	5983.621

The best model is adding nodes as a covariate, attaining an AIC of 5827.4. The second best model is differentiation of tumor and the third best is node4 (more than 4 positive lymph nodes). However, we should not include node4 since it is highly correlated with nodes. They are essentially measuring the same thing and in the dataset description it says it is unknown which of the two variables is correct, so this is a somewhat arbitrary selection.

```
colon.rx.nodes.differ = coxph(Surv(time,status)~rx + nodes + differ, colon.recur)
colon.rx.nodes.sex = coxph(Surv(time,status)~rx + nodes + sex, colon.recur)
colon.rx.nodes.age = coxph(Surv(time,status)~rx + nodes + age, colon.recur)
colon.rx.nodes.obstruct = coxph(Surv(time,status)~rx + nodes + obstruct, colon.recur)
colon.rx.nodes.perfor = coxph(Surv(time,status)~rx + nodes + perfor, colon.recur)
colon.rx.nodes.adhere = coxph(Surv(time,status)~rx + nodes + adhere, colon.recur)
colon.rx.nodes.extent = coxph(Surv(time,status)~rx + nodes + extent, colon.recur)
colon.rx.nodes.surg = coxph(Surv(time,status)~rx + nodes + surg, colon.recur)
```

```
AIC(colon.rx.nodes.differ, colon.rx.nodes.sex, colon.rx.nodes.age, colon.rx.nodes.obstruct, colon.rx.no
```

##		df	AIC
##	colon.rx.nodes.differ	4	5676.214
##	colon.rx.nodes.sex	4	5826.886
##	colon.rx.nodes.age	4	5828.648

```
## colon.rx.nodes.obstruct 4 5824.943
## colon.rx.nodes.perfor 4 5827.301
## colon.rx.nodes.adhere 4 5824.139
## colon.rx.nodes.extent 4 5802.637
## colon.rx.nodes.surg 4 5824.080
```

Adding differ as the next covariate is the best model, as our model now has an AIC of 5676.21

```
colon.rnd.sex = coxph(Surv(time,status)~rx + nodes + differ + sex, colon.recur)
colon.rnd.age = coxph(Surv(time,status)~rx + nodes + differ + age, colon.recur)
colon.rnd.obstruct = coxph(Surv(time,status)~rx + nodes + differ + obstruct, colon.recur)
colon.rnd.perfor = coxph(Surv(time,status)~rx + nodes + differ +perfor , colon.recur)
colon.rnd.adhere = coxph(Surv(time,status)~rx + nodes + differ +adhere , colon.recur)
colon.rnd.extent = coxph(Surv(time,status)~rx + nodes + differ + extent, colon.recur)
colon.rnd.surg = coxph(Surv(time,status)~rx + nodes + differ +surg , colon.recur)
AIC(colon.rnd.sex, colon.rnd.age, colon.rnd.obstruct, colon.rnd.perfor, colon.rnd.adhere, colon.rnd.extent)
```

```
##          df      AIC
## colon.rnd.sex      5 5675.069
## colon.rnd.age      5 5677.614
## colon.rnd.obstruct  5 5673.698
## colon.rnd.perfor   5 5676.290
## colon.rnd.adhere   5 5674.507
## colon.rnd.extent   5 5658.703
## colon.rnd.surg     5 5673.328
```

Adding extent as the next covariate is the best model, attaining an AIC of 5658.7

```
BIC(colon.rnd.extent,colon.rx.nodes.differ, colon.rx.nodes, colon.recur.cox)
```

```
##          df      BIC
## colon.rnd.extent      5 5679.205
## colon.rx.nodes.differ  4 5692.615
## colon.rx.nodes        3 5839.763
## colon.recur.cox       2 6068.505
```

Using BIC as a sanity check helps us to verify that adding extent as another covariate is still beneficial to our model's prediction accuracy while avoiding overfitting (too many covariates), attaining a BIC of 5679.21.

```
colon.rnde.sex = coxph(Surv(time,status)~rx + nodes + differ + extent + sex, colon.recur)
colon.rnde.age = coxph(Surv(time,status)~rx + nodes + differ + extent + age, colon.recur)
colon.rnde.obstruct = coxph(Surv(time,status)~rx + nodes + differ + extent + obstruct, colon.recur)
colon.rnde.perfor = coxph(Surv(time,status)~rx + nodes + differ + extent + perfor, colon.recur)
colon.rnde.adhere = coxph(Surv(time,status)~rx + nodes + differ + extent + adhere, colon.recur)
colon.rnde.surg = coxph(Surv(time,status)~rx + nodes + differ + extent + surg, colon.recur)

AIC(colon.rnde.sex, colon.rnde.age, colon.rnde.obstruct, colon.rnde.perfor,
    colon.rnde.adhere, colon.rnde.surg)
```

```
##          df      AIC
## colon.rnde.sex      6 5657.874
## colon.rnde.age      6 5660.177
## colon.rnde.obstruct  6 5657.612
## colon.rnde.perfor   6 5659.534
## colon.rnde.adhere   6 5658.749
## colon.rnde.surg     6 5655.546
```

```
BIC(colon.rnde.sex, colon.rnde.age, colon.rnde.obstruct, colon.rnde.perfor,
    colon.rnde.adhere, colon.rnde.surg)
```

```
##                df      BIC
## colon.rnde.sex    6 5682.476
## colon.rnde.age    6 5684.779
## colon.rnde.obstruct 6 5682.214
## colon.rnde.perfor  6 5684.136
## colon.rnde.adhere  6 5683.351
## colon.rnde.surg    6 5680.148
```

Adding Surg (time from surgery until registration into the study) as our next covariate again lowers our AIC to 5655.55

```
colon.rndes.sex = coxph(Surv(time,status)~rx + nodes + differ + extent + surg + sex, colon.recur)
colon.rndes.age = coxph(Surv(time,status)~rx + nodes + differ + extent + surg + age, colon.recur)
colon.rndes.obstruct = coxph(Surv(time,status)~rx + nodes + differ + extent + surg + obstruct, colon.recur)
colon.rndes.perfor = coxph(Surv(time,status)~rx + nodes + differ + extent + surg + perfor, colon.recur)
colon.rndes.adhere = coxph(Surv(time,status)~rx + nodes + differ + extent + surg + adhere, colon.recur)
```

```
AIC(colon.rndes.sex, colon.rndes.age, colon.rndes.obstruct, colon.rndes.perfor, colon.rndes.adhere)
```

```
##                df      AIC
## colon.rndes.sex    7 5654.365
## colon.rndes.age    7 5656.926
## colon.rndes.obstruct 7 5654.744
## colon.rndes.perfor  7 5656.414
## colon.rndes.adhere  7 5655.639
```

We attain a slightly lower AIC of 5654.37 by adding sex as our next covariate. However, since the change is small we will verify this using BIC score too.

```
BIC(colon.rndes.sex, colon.rnde.surg)
```

```
##                df      BIC
## colon.rndes.sex    7 5683.067
## colon.rnde.surg    6 5680.148
```

Using BIC score instead, which penalizes more covariates more intensely than AIC, we see that our previous model is actually preferable, so we will stop adding covariates at Surg.

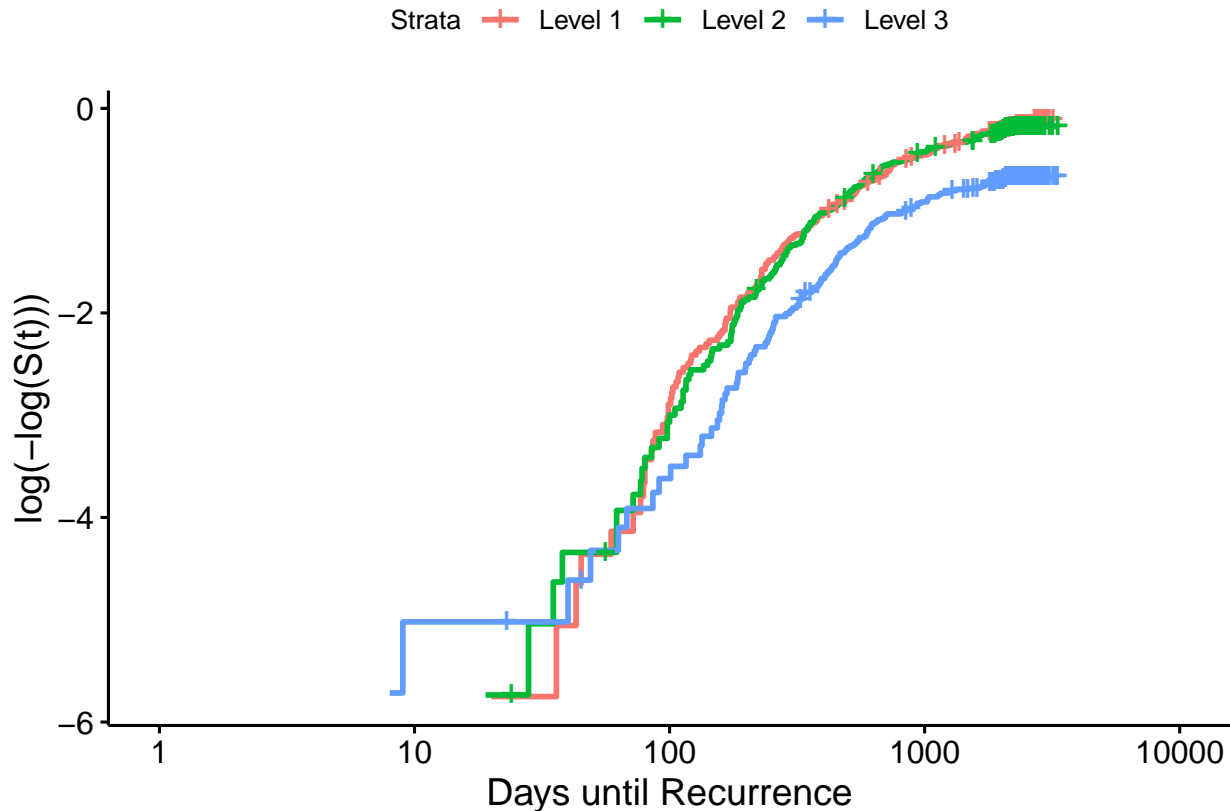
```
anova(colon.rndes.sex)
```

```
## Analysis of Deviance Table
## Cox model: response is Surv(time, status)
## Terms added sequentially (first to last)
##
##          loglik   Chisq Df Pr(>|Chi|)
## NULL      -2877.1
## rx        -2865.6 23.1505  2  9.396e-06 ***
## nodes     -2835.8 59.6074  1  1.158e-14 ***
## differ    -2834.1  3.2934  1   0.06956 .
## extent    -2824.3 19.5108  1  1.000e-05 ***
## surg       -2821.8  5.1569  1   0.02315 *
## sex        -2820.2  3.1816  1   0.07447 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


We also verify this by checking that if we add sex as a covariate, it is not significantly affecting time until recurrence from the likelihood ratio test.

Checking Proportional Hazards Assumptions

```
colon.fit = survfit(Surv(time, status) ~ rx, colon.recur)
ggsurvplot(colon.fit, colon.recur,
            legend.labs=c('Level 1', 'Level 2', 'Level 3'),
            fun='cloglog') +
  labs(x='Days until Recurrence')
```



```
cox.zph(colon.recur.cox)
```

```
##      chisq df    p
## rx      0.301  2 0.86
## GLOBAL 0.301  2 0.86
```

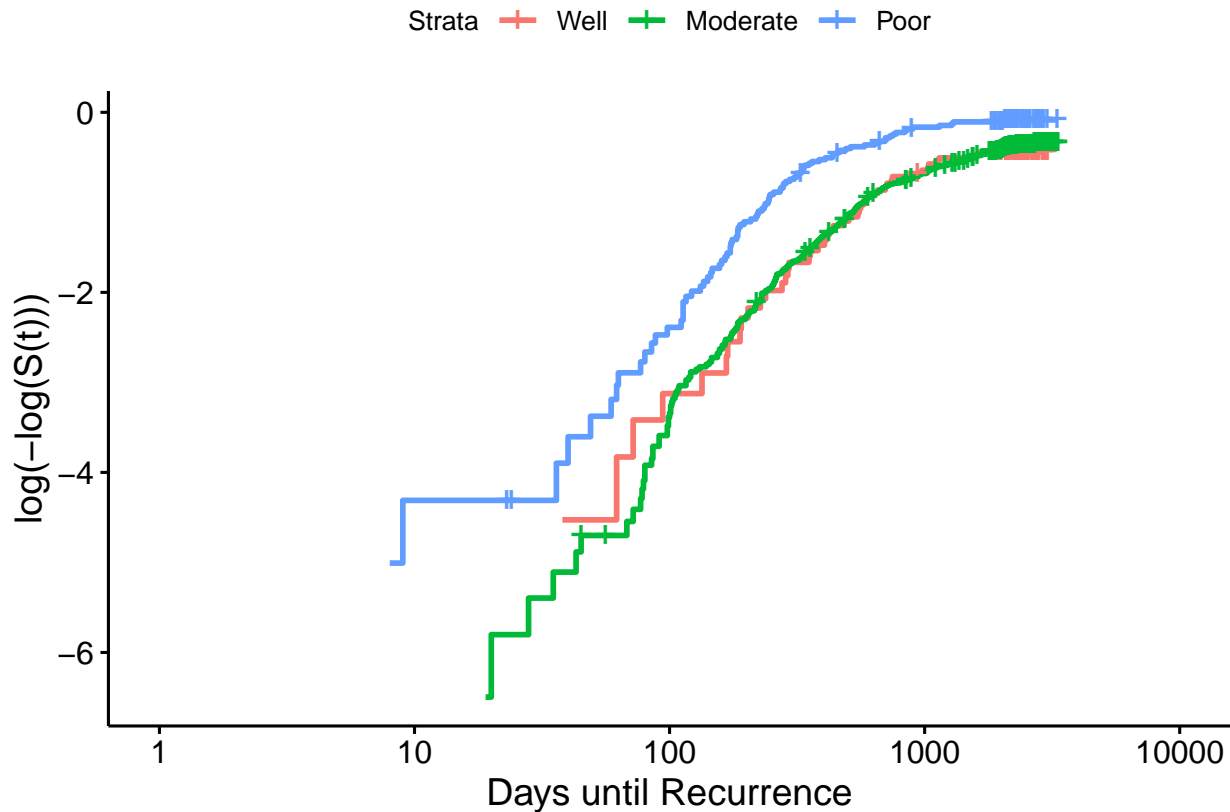
The proportional hazards assumption seems violated from the log-log plot for Baseline and low toxicity treatment, but we have already observed that those two treatments do not significantly differ in hazard rates. It appears that the difference between low toxicity and medium toxicity does not violate the proportional hazards assumption, and we verify this using the Schoenfeld residuals test attaining a p-value of 0.86 for the rx covariate.

```
cox.zph(colon.rnde.surg)
```

```
##      chisq df      p
## rx      0.454  2 0.79696
## nodes   1.295  1 0.25504
## differ 13.525  1 0.00024
## extent  0.140  1 0.70802
```

```
## surg      1.827  1 0.17644
## GLOBAL 16.112  6 0.01316
```

```
colon.fit.differ = survfit(Surv(time, status) ~ differ, colon.recur)
ggsurvplot(colon.fit.differ, colon.recur,
  legend.labs=c('Well', 'Moderate', 'Poor'), fun='cloglog') +
  labs(x='Days until Recurrence')
```



From our cox ZPH test we see that the differentiation of tumor covariate violates the proportional hazards assumption with a p-value of 0.00024 and we verify this using a log-log plot. We can observe many cross-over points between the plot for 'moderate' differentiation and 'well' differentiation. To account for this violation, we will attempt to stratify on the differentiation covariate and produce a stratified Cox PH model.

```
colon.rnde.surg = coxph(Surv(time,status)~rx + nodes + strata(differ) + extent + surg, colon.recur)
anova(colon.rnde.surg)
```

```
## Analysis of Deviance Table
## Cox model: response is Surv(time, status)
## Terms added sequentially (first to last)
##
##          loglik   Chisq Df Pr(>|Chi|)
## NULL      -2529.3
## rx        -2517.6 23.3621  2  8.453e-06 ***
## nodes     -2491.5 52.1831  1  5.056e-13 ***
## extent    -2481.4 20.1808  1  7.046e-06 ***
## surg      -2478.7  5.4127  1  0.01999 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cox.zph(colon.rnde.surg)
```

```
##           chisq df    p
## rx      0.76738  2 0.68
## nodes   0.25978  1 0.61
## extent  0.00222  1 0.96
## surg    1.70609  1 0.19
## GLOBAL  2.61561  5 0.76
```

```
mean(colon.recur$nodes, na.rm=TRUE)
```

```
## [1] 3.659715
```

```
max(colon.recur$nodes, na.rm=TRUE)
```

```
## [1] 33
```

Our AIC drops to 4967.46 after stratifying on the differentiation variable instead of simply including it as a covariate in our model. This is a significant decrease in AIC most likely due to the fact that differentiation should have never been included as a covariate in the model, as it violated the proportional hazards assumption.

To verify again that none of our covariates violate this assumption, we will run the Cox ZPH test on our new stratified model. We observe that all of our p-values are greater than 0.05 and we can keep them in our model as covariates.

We verify again post-stratification that all of our covariates are significant. We can begin our model interpretation as follows:

- After accounting for other covariate effects, the impact on hazard rate of the Levamisole+5FU treatment remains a factor of 0.6.
- For each additional lymph nodes with detectable cancer that is observed at the start of the study, the hazard rate for that individual increases by a factor of 1.078. This is not a huge effect, but it is important to note that the number of lymph nodes with cancer were as high as 33 in our data set, so the hazard rate could gradually increase to significantly high levels as this covariate increases for certain subjects. However, the mean of this covariate is only ~ 3.7 nodes
- For our 4 levels of increasing local spread extent of the colon cancer, hazard rate increases by an alarming factor of 1.66. This is an intuitive finding, as the extent of cancer spread is a strong indicator of a patient having a recurrence or death.
- If time from surgery to registration is classified as 'long', hazard rate is increased by a factor of 1.28. This is an interesting finding and indicates that the time spent between surgery and registration is vital to the time it took patients to have a disease recurrence or death. This is also an intuitive finding, as a patient that had a long time in between is experiencing more time that could contribute to the growth of cancer spread.

Death Subset

Using the same exact process as above for the data with etype=1 (recurrence of cancer), we find that for etype = 2 (death), the model had the same covariates until adding the 6th covariate. Adding obstruct to the death cox model gave us the next best model which is different from the recur data, where at this point, adding sex as our next covariate gave us the lower AIC score. However, in both subsets of data, adding the 6th covariate gave us a higher BIC score, allowing us to leave it out and have the same covariates (rx, node, differ, extent, surg).

The best model in the data for patients with a recorded death adds nodes with an AIC of 5620.639. The second best model adds differ with an AIC of 5678.121 and the third best model adds node4 with an AIC of 5764.921.

After adding differ we got an AIC of 5458.958.

Adding extent give us an AIC of 5441.828.

Adding surg gives an AIC of 5438.274.

Adding obstruct gives us an AIC of 5436.907, which is different from the recur data, where at this point, adding sex as our next covariate gave us the lower AIC score.

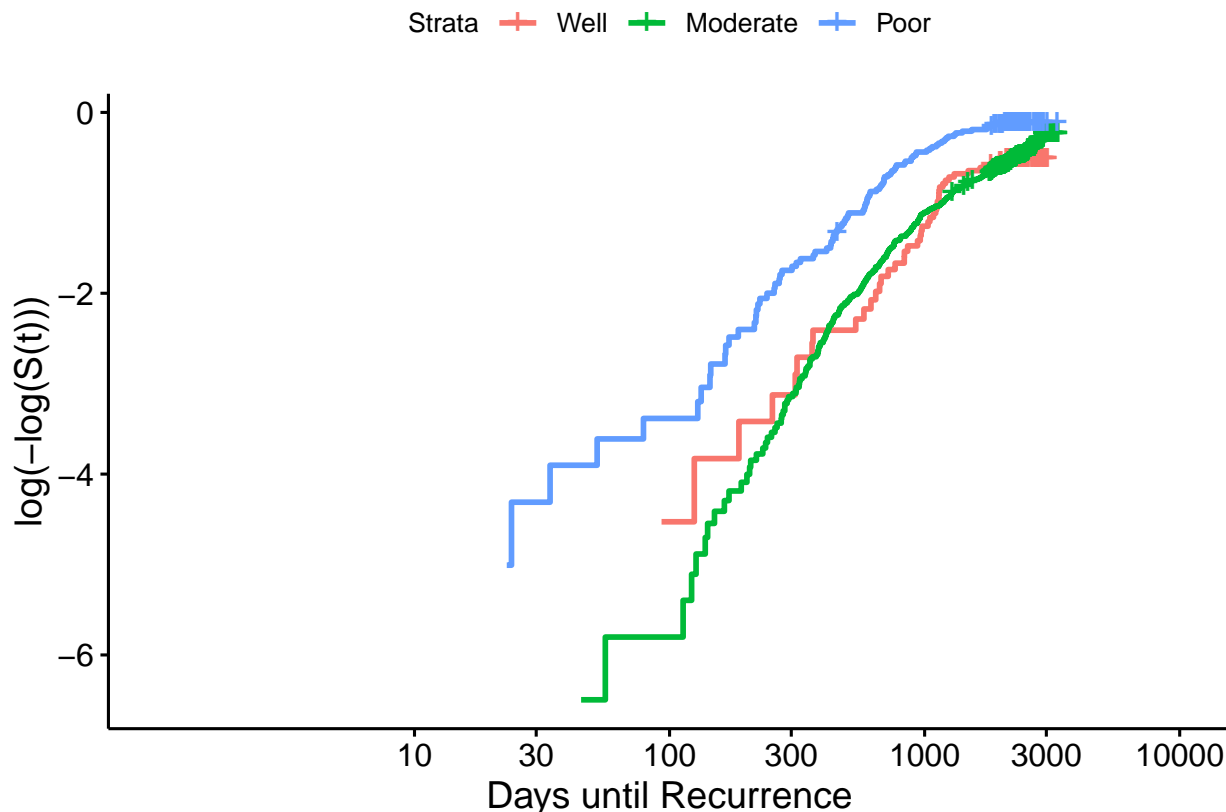
By using BIC, which penalizes the number of covariates more than AIC, we can see that the previous model, without obstruct, is a better model as it gives a BIC of 5462.657. Therefore, we will use the exact same covariates as the recurrence model.

Checking Proportional Hazards Assumptions

```
cox.zph(death.rnde.surg)
```

```
##          chisq df      p
## rx          2.6512  2 0.26565
## nodes        0.0206  1 0.88595
## differ     12.3575  1 0.00044
## extent       3.5707  1 0.05881
## surg         0.0213  1 0.88386
## GLOBAL     18.4058  6 0.00529
```

```
colon.fit.differ = survfit(Surv(time, status) ~ differ, colon.death)
ggsurvplot(colon.fit.differ, colon.death, legend.labs=c('Well', 'Moderate', 'Poor'), fun='cloglog') +
  labs(x='Days until Recurrence')
```



We can also observe, that similarly to our recurrence model the differentiation covariate again violates the PH assumption. We can verify this using a log-log plot, which shows many cross-over points between 'moderate'

differentiation and ‘well’ differentiation. We will remedy this using a stratification on the differentiation variable.

```
death.rnde.surg <- coxph(Surv(time,status)~rx + nodes + strata(differ) + extent + surg, colon.death)
AIC(death.rnde.surg)
```

```
## [1] 4769.913
```

Our AIC drops to 4769.91 after stratifying on the differentiation covariate instead of simply using it as a covariate in our model. This is a significant decrease and proves that our previous model had an alarming violation of the PH assumption.

```
cox.zph(death.rnde.surg)
```

```
##          chisq df    p
## rx          2.2824 2 0.32
## nodes       0.7298 1 0.39
## extent     2.6415 1 0.10
## surg        0.0332 1 0.86
## GLOBAL     6.2753 5 0.28
```

We run cox ZPH again on this new stratified cox PH model to ensure no violations and observe that we have successfully remedied our issue. All of our p-values from the cox ZPH test are above 0.05.

```
anova(death.rnde.surg)
```

```
## Analysis of Deviance Table
## Cox model: response is Surv(time, status)
## Terms added sequentially (first to last)
##
##          loglik    Chisq Df Pr(>|Chi|)
## NULL        -2429.0
## rx          -2423.3 11.2821  2    0.003549 **
## nodes       -2393.0 60.6502  1    6.817e-15 ***
## extent     -2382.8 20.4394  1    6.155e-06 ***
## surg        -2380.0  5.6674  1    0.017283 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(death.rnde.surg)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ rx + nodes + strata(differ) +
##       extent + surg, data = colon.death)
##
##      n= 888, number of events= 430
##      (41 observations deleted due to missingness)
##
##              coef exp(coef)  se(coef)      z Pr(>|z|)
## rxLev        -0.081997  0.921275  0.114478 -0.716  0.47382
## rxLev+5FU    -0.364352  0.694646  0.121613 -2.996  0.00274 **
## nodes         0.084433  1.088100  0.009429  8.955 < 2e-16 ***
## extent       0.506396  1.659300  0.116614  4.342 1.41e-05 ***
## surg         0.257217  1.293326  0.106194  2.422  0.01543 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
```

```

## rxLev      0.9213      1.0855      0.7361      1.1530
## rxLev+5FU  0.6946      1.4396      0.5473      0.8816
## nodes      1.0881      0.9190      1.0682      1.1084
## extent     1.6593      0.6027      1.3203      2.0854
## surg       1.2933      0.7732      1.0503      1.5926
##
## Concordance= 0.657 (se = 0.015 )
## Likelihood ratio test= 98.04 on 5 df, p=<2e-16
## Wald test           = 115.7 on 5 df, p=<2e-16
## Score (logrank) test = 119.5 on 5 df, p=<2e-16

```

We verify again post-stratification that all of our covariates are significant. We can begin our model interpretation as follows:

- After accounting for other covariate effects, the impact on hazard rate of the Levamisole+5FU treatment is slightly increased to a factor of 0.694. In other words, its influence on decreasing hazard rate is slightly lessened after accounting for other variables.
- For each additional lymph nodes with detectable cancer that is observed at the start of the study, the hazard rate for that individual increases by a factor of 1.088
- For our 4 levels of increasing local spread extent of the colon cancer, hazard rate increases by an alarming factor of 1.66.
- If time from surgery to registration is classified as 'long', hazard rate is increased by a factor of 1.29.

3. Advanced Models

Recurrent Events Model

```
colon$start=0
colon$stop = colon$time

colon.cox = coxph(Surv(start,stop,status)~rx + strata(etype), colon)
summary(colon.cox)
```

```
## Call:
## coxph(formula = Surv(start, stop, status) ~ rx + strata(etype),
##       data = colon)
##
##      n= 1858, number of events= 920
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## rxLev          -0.02069   0.97952  0.07683 -0.269    0.788
## rxLev+5FU      -0.44275   0.64227  0.08391 -5.276 1.32e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## rxLev              0.9795        1.021    0.8426    1.1387
## rxLev+5FU          0.6423        1.557    0.5449    0.7571
##
## Concordance= 0.545 (se = 0.009 )
## Likelihood ratio test= 35.54 on 2 df,  p=2e-08
## Wald test               = 33.39 on 2 df,  p=6e-08
## Score (logrank) test = 33.91 on 2 df,  p=4e-08
anova(colon.cox)
```

```
## Analysis of Deviance Table
## Cox model: response is Surv(start, stop, status)
## Terms added sequentially (first to last)
##
##      loglik  Chisq Df Pr(>|Chi|)
## NULL -5970.5
## rx   -5952.7 35.536  2   1.92e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

To produce a combined model that considers both recurrence and death, we will restructure our data into stop-start data format. We will use the marginal recurrent event approach, since each event is of a different nature. The first occurrence for each subject is a recurrence of cancer, while the second occurrence is death. Therefore, each start time will be 0 and each stop time will be time of death or time of recurrence.

In our recurrent events model, we again observe with a p-value of 0.788 that the Levamisole treatment in itself is not effective in reducing time until recurrence or death. However, Levamisole treatment combined with the moderately toxic 5FU treatment is effective in reducing time until recurrence or death, as it has a p-value of ~ 0 . Similarly to our recurrence and death analyses which were conducted separately, we reach the same conclusion that is noted by the researchers of the original study (Laurie 1989).

```
colon.CI = exp(confint(colon.cox))
colon.CI
```

```
##           2.5 %    97.5 %
## rxLev      0.8425925 1.1387073
## rxLev+5FU 0.5448655 0.7570817
```

We confirm with 95% confidence that Levamisole treatment is not effective in itself using a confidence interval [0.843, 1.139] that contains 1. Our confidence interval for Levamisole+5FU treatment is quite small [0.545, 0.757] and does not contain 1. We confirm our previous conclusion with 95% confidence that the combined treatment is effective in reducing time until recurrence or death.

```
exp(coef(colon.cox))
```

```
##      rxLev rxLev+5FU
## 0.9795235 0.6422676
```

The instantaneous hazard rate for Levamisole+5FU treatment is 0.642 times as large as the baseline hazard rate. We conclude that the treatment is indeed effective in significantly reducing time until a recurrence or death. This is the same conclusion that was derived from the recurrence data set as well as the death data set, but we can confirm that the combination treatment of low-toxicity and medium-toxicity compounds retains its effectiveness in the combined model of both event types.

Adding Covariates to the Combined Data Set

Again using forward stepwise selection, AIC scores, and verifying AIC scores using BIC when improvements are minimal, we will add the other covariates into our full model.

Similar to our data subsets where etype = 1 and etype = 2, adding nodes as the next covariate after rx gave us the lowest AIC, 11443.35. Node4 follows with the next lowest AIC score of 11743.73. We will not use node4 as a covariate since it is essentially measuring the same quantity as the nodes covariate. We do not want to use highly correlated covariates in our cox PH model.

Adding differ gives us the lowest AIC of 11128.39

Adding extent as our next covariate gives us an AIC of 11091.87

Adding surg gives us the lowest AIC of 11083.24

Adding obstruct gives us an AIC of 11079.07. This change is minimal, so we will verify it using BIC.

```
BIC(full.rnde.surg, full.rndes.obstruct)
```

```
##           df      BIC
## full.rnde.surg      6 11111.89
## full.rndes.obstruct  7 11112.50
```

By using BIC, which penalizes the number of covariates more strictly than AIC, we can see that the model with surg is a better model as it gives a BIC of 11111.89 as opposed to 11112.5 for the model that adds obstruct as another covariate. We will stop adding covariates here.

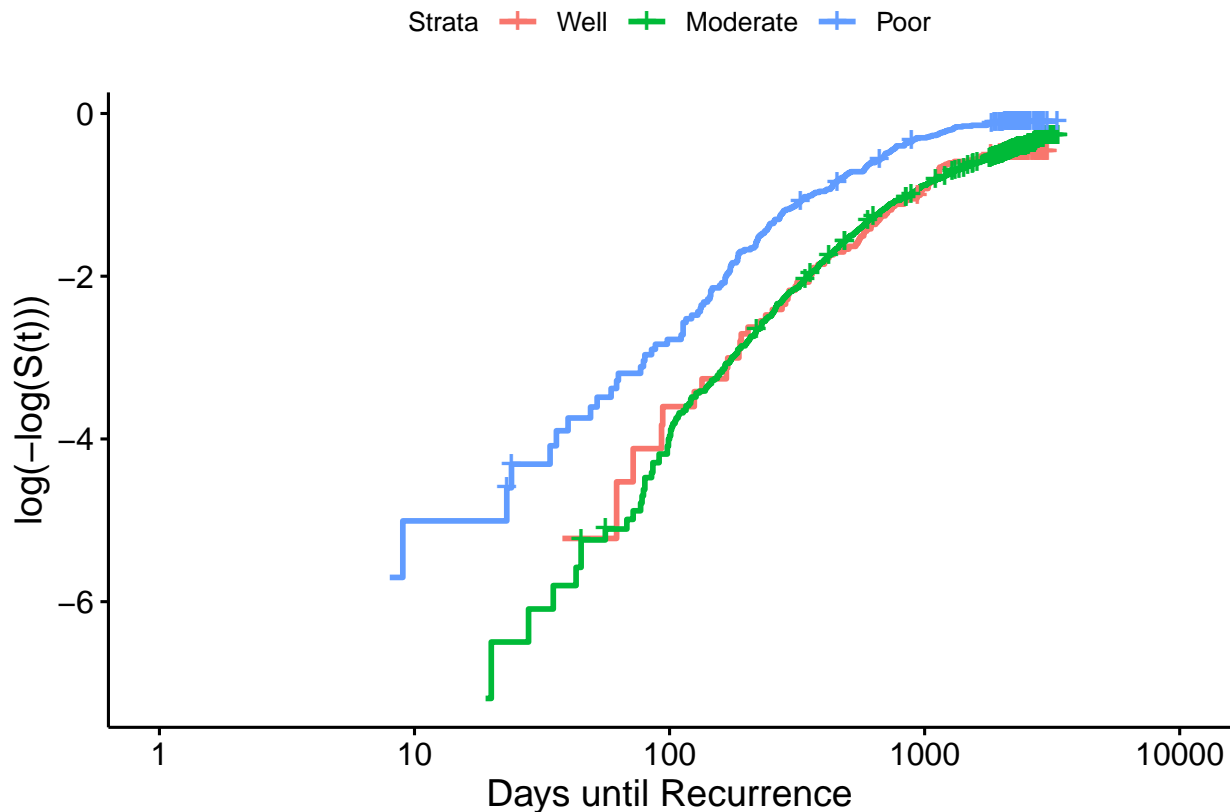
We verify using the likelihood ratio test that all of our covariates have a significant effect on time until recurrence or death.

```
cox.zph(full.rnde.surg)
```

```
##      chisq df      p
## rx      1.3046 2 0.52085
## nodes   0.0526 1 0.81855
## differ 21.4175 1 3.7e-06
## extent  2.6242 1 0.10525
## surg    0.5349 1 0.46454
## GLOBAL 25.2472 6 0.00031
```



```
colon.fit.differ = survfit(Surv(start, stop, status) ~ differ, colon)
ggsurvplot(colon.fit.differ, colon, legend.labs=c('Well', 'Moderate', 'Poor'), fun='cloglog') +
  labs(x='Days until Recurrence')
```



Using a cox ZPH test and a log-log plot, we can see that although differ was previously determined to have a significant effect, it violates the proportional hazards assumption. To fix this, we will construct a stratified model that stratifies on the 3 levels of tumor differentiation.

```
comb.strat <- coxph(Surv(start, stop, status) ~ rx + nodes + strata(differ) + extent + surg + strata(et,
cox.zph(comb.strat)
```

```
##          chisq df    p
## rx          1.341  2 0.51
## nodes        0.346  1 0.56
## extent       1.496  1 0.22
## surg         0.458  1 0.50
## GLOBAL       3.907  5 0.56
```

```
AIC(comb.strat)
```

```
## [1] 9728.766
```

We verify again with the cox ZPH test that none of our covariates are in violation of the proportional hazards assumption after stratifying on tumor differentiation. In addition, we have massively reduced our AIC score to 9728.77 by fixing this violation.

```
anova(comb.strat)
```

```
## Analysis of Deviance Table
## Cox model: response is Surv(start, stop, status)
```

```
## Terms added sequentially (first to last)
##
##      loglik   Chisq Df Pr(>|Chi|)
## NULL      -4958.3
## rx        -4941.5  33.637  2  4.963e-08 ***
## nodes     -4885.1 112.637  1 < 2.2e-16 ***
## extent    -4864.9  40.504  1  1.962e-10 ***
## surg      -4859.4  11.009  1  0.0009067 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(comb.strat)

## Call:
## coxph(formula = Surv(start, stop, status) ~ rx + nodes + strata(differ) +
##      extent + surg + strata(etype), data = colon)
##
##      n= 1776, number of events= 876
##      (82 observations deleted due to missingness)
##
##      coef exp(coef) se(coef)      z Pr(>|z|)
## rxLev     -0.075715  0.927080  0.079711 -0.950 0.342174
## rxLev+5FU -0.436817  0.646090  0.085955 -5.082 3.74e-07 ***
## nodes      0.079517  1.082764  0.006623 12.006 < 2e-16 ***
## extent     0.505874  1.658435  0.082628  6.122 9.22e-10 ***
## surg       0.250680  1.284899  0.074281  3.375 0.000739 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##      exp(coef) exp(-coef) lower .95 upper .95
## rxLev      0.9271      1.0787      0.7930      1.0838
## rxLev+5FU   0.6461      1.5478      0.5459      0.7646
## nodes       1.0828      0.9236      1.0688      1.0969
## extent      1.6584      0.6030      1.4105      1.9500
## surg        1.2849      0.7783      1.1108      1.4863
##
## Concordance= 0.655 (se = 0.01 )
## Likelihood ratio test= 197.8 on 5 df, p=<2e-16
## Wald test              = 228.1 on 5 df, p=<2e-16
## Score (logrank) test = 235.6 on 5 df, p=<2e-16
```

We verify again post-stratification that all of our covariates are significant. We can begin our model interpretation as follows:

- After accounting for other covariate effects, the impact on hazard rate of the Levamisole+5FU treatment is slightly increased to a factor of 0.646. In other words, its influence on decreasing hazard rate is slightly lessened after accounting for other variables.
- For each additional lymph nodes with detectable cancer that is observed at the start of the study, the hazard rate for that individual increases by a factor of 1.083
- For our 4 levels of increasing local spread extent of the colon cancer, hazard rate increases by an alarming factor of 1.66.
- If time from surgery to registration is classified as 'long', hazard rate is increased by a factor of 1.28.

Conclusion

Clearly, we have reached some interesting conclusions in our recurrence subset, death subset, and overall death & recurrence analysis. The covariate effects seem to be quite similar across all of these analyses and influence time until recurrence and death in similar ways. One surprise is that our combined model, which stratifies on event type and uses a marginal time approach, had very similar conclusions to our recurrence and death subset analyses. We can say with confidence that overall, low toxicity Levamisole is not an effective treatment in itself to treat colon cancer and reduce time until a recurrence or patient death. It must be combined with 5FU moderately toxic treatment to have a significant effect. The other measurements that were included in our model and discussed throughout our analyses are important indicators for researchers to predict which patients will have a higher rate of recurrence or death. Therefore, it is important when tracking patients to not only measure the treatment given, but the state of the disease itself at study onset. Measurements such as # nodes with cancer, extent of local spread, and time from surgery to registration are absolutely vital to make accurate predictions in addition to measuring the treatment given to a particular patient. If a patient has high local spread, many lymph nodes with detectable cancer, and a long time from surgery to registration, they can reasonably be classified as a 'high risk' for recurrence or death individual.

4. Citations

Laurie, J A et al. "Surgical adjuvant therapy of large-bowel carcinoma: an evaluation of levamisole and the combination of levamisole and fluorouracil. The North Central Cancer Treatment Group and the Mayo Clinic." *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* vol. 7,10 (1989): 1447-56. doi:10.1200/JCO.1989.7.10.1447