Survival Analysis of Recurrence of Adjuvant Chemotherapy for Colon Cancer

Kaitlyn Wang, Yimiao Pang

2022-12-04

1. Data Tidy

In this section, we subset the colon data to only include the event of recurrence.

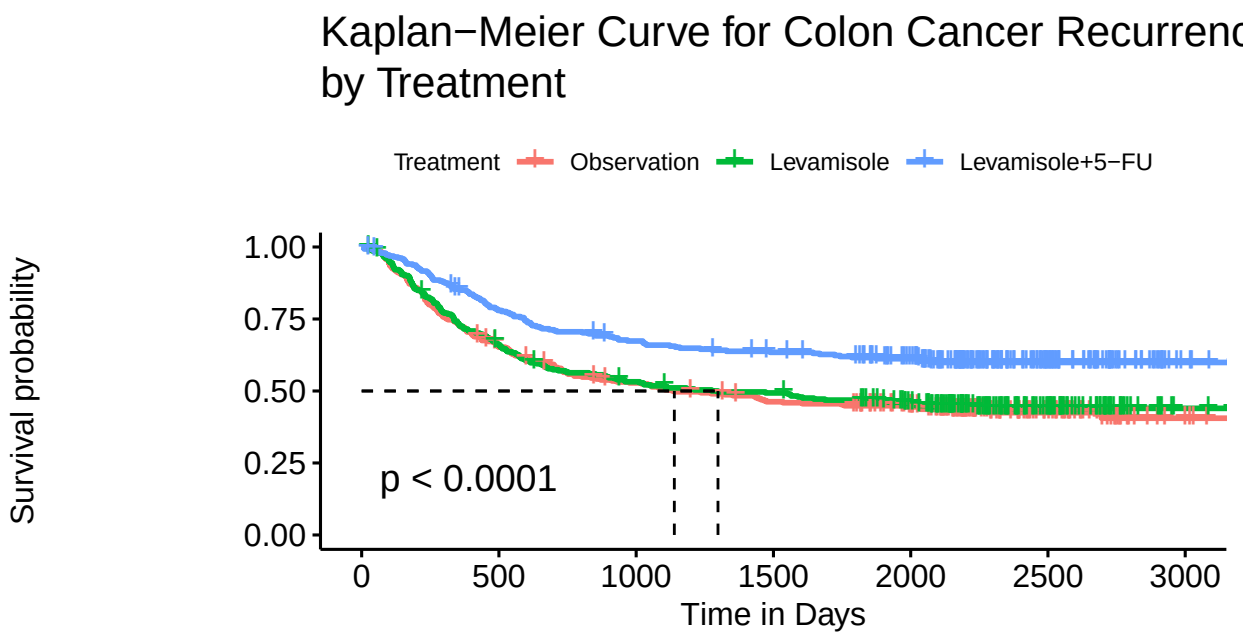
```
# import the dataset from survival package
colon <- survival::colon

# recurrence dataset
colon.recur <- colon %>% filter(etype == '1') %>%
               select(-c(id,study,etype, node4)) %>%
               drop_na() %>%
               mutate(rx = as.factor(rx))
```

2. Kaplan-Meier Survival Estimate

```
recur.fit <- survfit(Surv(time, status) ~ rx, data = colon.recur)

ggsurvplot(recur.fit, conf.int = F, break.time.by = 500, pval = TRUE,
           font.x.size = 12, font.y.size = 12, font.legend.size = 9, surv.median.line = "hv",
           legend.title = "Treatment", legend.labs = c("Observation", "Levamisole","Levamisole+5-FU"),
           title = "Kaplan-Meier Curve for Colon Cancer Recurrence \nby Treatment",
           xlab = "Time in Days",
           risk.table = T, risk.table.height = 0.25, risk.table.fontsize = 4,
           tables.theme = theme_cleantable())
```



Number at risk

Observation	305	196	155	133	112	39	5
Levamisole	294	191	152	140	117	47	4
Levamisole+5−FU	289	221	189	175	149	60	7

From the Kaplan-Meier curve shown above, there is significant difference among the three treatment groups given p-value < 0.0001, and patients treated with levamisole + 5-FU have higher survival probability than patients with no further treatment and patients who received the treatment with levamisole alone. The median survival time for observation group and levamisole group are approximately 1100 days and 1300 days, respectively. However, until the end of the trial, the survival probability of Levamisole + 5-FU treatment group is greater than 50% as we fail to observe the curve crossing the 50% line.

3. Log-Rank Test

Given the difference of survival probability among the three treatment groups, we conduct a Log-rank hypothesis test to test the null hypothesis of no difference among the three treatments in the mortality model.

```
survdiff(Surv(time, status) ~ rx, data = colon.recur)
```

```
## Call:
## survdiff(formula = Surv(time, status) ~ rx, data = colon.recur)
##
##                N Observed Expected (O-E)^2/E (O-E)^2/V
## rx=Obs       305      172      145      5.13      7.60
## rx=Lev       294      161      141      2.84      4.15
## rx=Lev+5FU 289      113      160     13.93     21.80
##
##  Chisq= 22  on 2 degrees of freedom, p= 2e-05
```

From this log-rank test, we observed a p-value equal 0.00002, which indicates a significant difference among treatment groups at a 0.05 level.

4. Cox PH Model

4.1 Model Selection

We now conduct automatic stepwise selection with Akaike information criterion (AIC) to determine the covariates that best represent an appropriate cox proportional hazards model for the event of recurrence.

```
# full model
r.model.full <- coxph(Surv(time, status) ~ ., data = colon.recur)

# stepwise selection with AIC criterion
r.model.aic <- step(r.model.full, direction = "both", k = 2)
```

```
## Start:  AIC=5656.88
## Surv(time, status) ~ rx + sex + age + obstruct + perfor + adhere +
##     nodes + differ + extent + surg
##
##            Df    AIC
## - perfor    1 5655.3
## - age       1 5655.5
## - adhere    1 5656.4
## <none>        5656.9
## - obstruct  1 5657.0
## - differ    1 5657.0
## - sex       1 5658.0
## - surg      1 5660.2
## - extent    1 5671.1
## - rx        2 5673.7
## - nodes     1 5706.6
##
## Step:  AIC=5655.3
## Surv(time, status) ~ rx + sex + age + obstruct + adhere + nodes +
##     differ + extent + surg
##
##            Df    AIC
## - age       1 5654.0
## - adhere    1 5655.2
## <none>        5655.3
## - differ    1 5655.5
## - obstruct  1 5655.7
## - sex       1 5656.3
## + perfor    1 5656.9
## - surg      1 5658.6
## - extent    1 5669.8
## - rx        2 5671.9
## - nodes     1 5704.9
##
## Step:  AIC=5654.01
## Surv(time, status) ~ rx + sex + obstruct + adhere + nodes + differ +
```

```
##     extent + surg
##
##            Df    AIC
## - adhere    1 5653.7
## <none>        5654.0
## - differ    1 5654.2
## - obstruct  1 5654.7
## - sex       1 5654.9
## + age       1 5655.3
## + perfor    1 5655.5
## - surg      1 5657.2
## - extent    1 5668.7
## - rx        2 5671.0
## - nodes     1 5704.7
##
## Step:  AIC=5653.72
## Surv(time, status) ~ rx + sex + obstruct + nodes + differ + extent +
##     surg
##
##            Df    AIC
## <none>        5653.7
## + adhere    1 5654.0
## - differ    1 5654.3
## - obstruct  1 5654.4
## - sex       1 5654.7
## + perfor    1 5654.8
## + age       1 5655.2
## - surg      1 5657.0
## - extent    1 5669.8
## - rx        2 5670.9
## - nodes     1 5703.8
```

The selected model with stepwise selection method is: Surv(time, status) ~ differ + obstruct + sex + surg + extent + rx + nodes

Then we use the ANOVA procedure to confirm if each of the covariates selected by the stepwise selection method is significant to be included in the Cox Proportional Model.

```
anova(r.model.aic)
```

```
## 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 ***
## sex      -2864.6  1.9131  1    0.16662
## obstruct -2863.1  3.0451  1    0.08098 .
## nodes    -2832.3 61.5151  1  4.394e-15 ***
## differ   -2830.4  3.8673  1    0.04924 *
## extent   -2821.5 17.8128  1  2.437e-05 ***
## surg     -2818.9  5.2438  1    0.02203 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-values for covariates “rx”, “nodes”, “differ”, “extent”, and “surg” are less than 0.05, except for “sex” and “obstruct”, indicating that they have a significant effect on time until recurrence. Thus, we will include these 5 covariates in our Cox PH model.

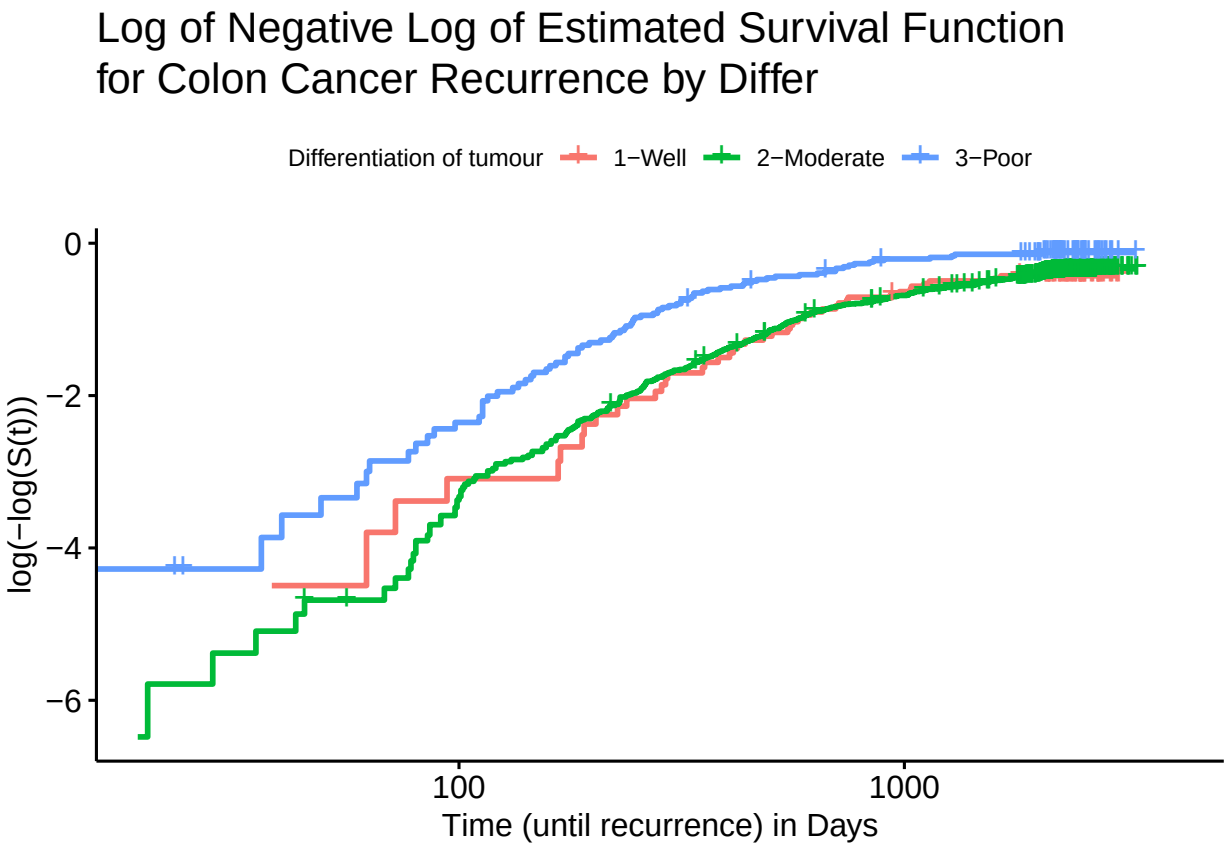
Therefore, we obtain the resulting model, which is Surv(time, status) ~ rx + differ + extent + surg + nodes.

4.2 Model Diagnostic

4.2.1 Check proportionality of hazard ratios

Log of Negative Log of Estimated Survival Function

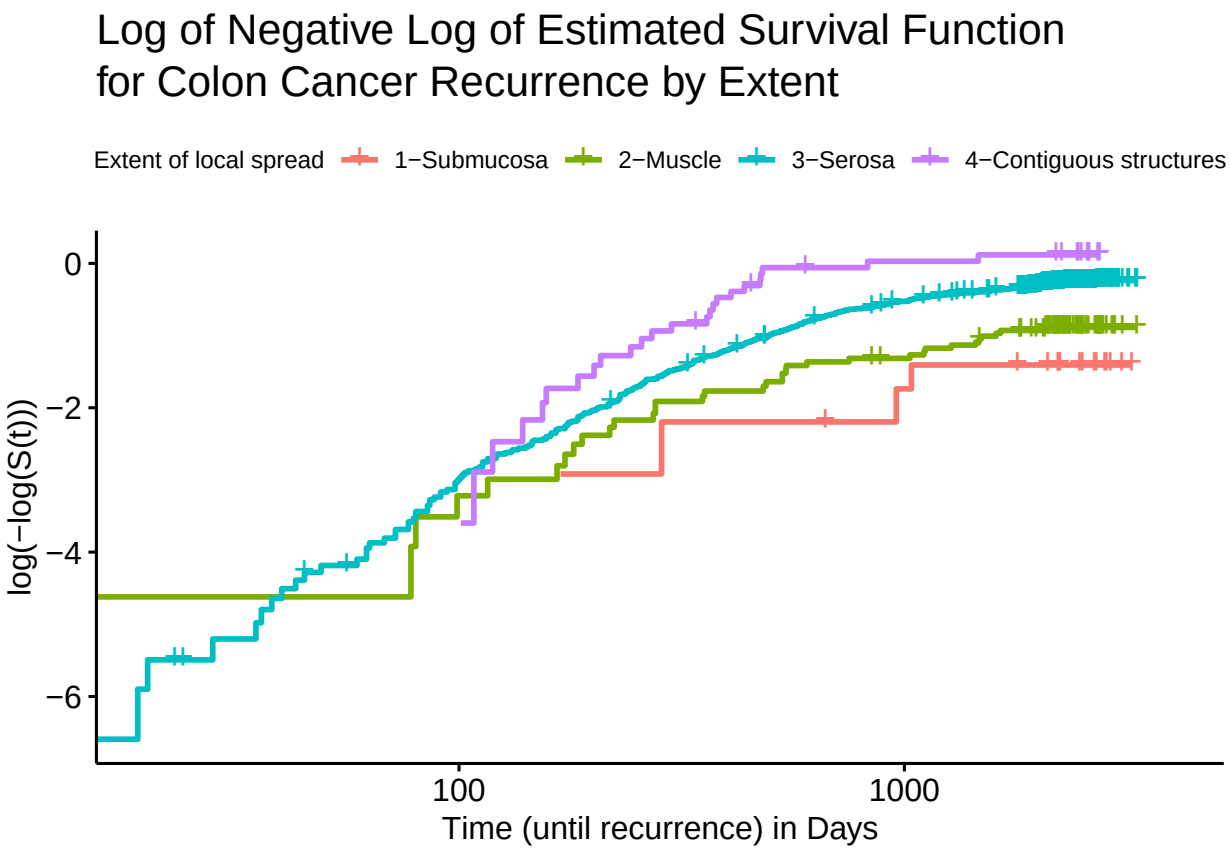
```
r.differ.fit <- survfit(Surv(time, status) ~ differ, data = colon.recur)
cll.differ = ggsurvplot(r.differ.fit, conf.int = F, font.x.size = 12, font.y.size = 12, font.legend.size
           fun = "cloglog",
           xlim = c(20, 4000),
           xlab = "Time (until recurrence) in Days",
           legend.lab = c("1-Well","2-Moderate","3-Poor"),
           legend.title = "Differentiation of tumour",
           title = "Log of Negative Log of Estimated Survival Function \nfor Colon Cancer Recurrence by
cll.differ
```



According to the plot, the curves in the C-log-log plot are crossing over after 100 days. The assumption of proportional hazard among the three differ groups is violated as we fail to see three parallel lines against log time.

We continue to plot the C-log-log plot for the covariate extent.

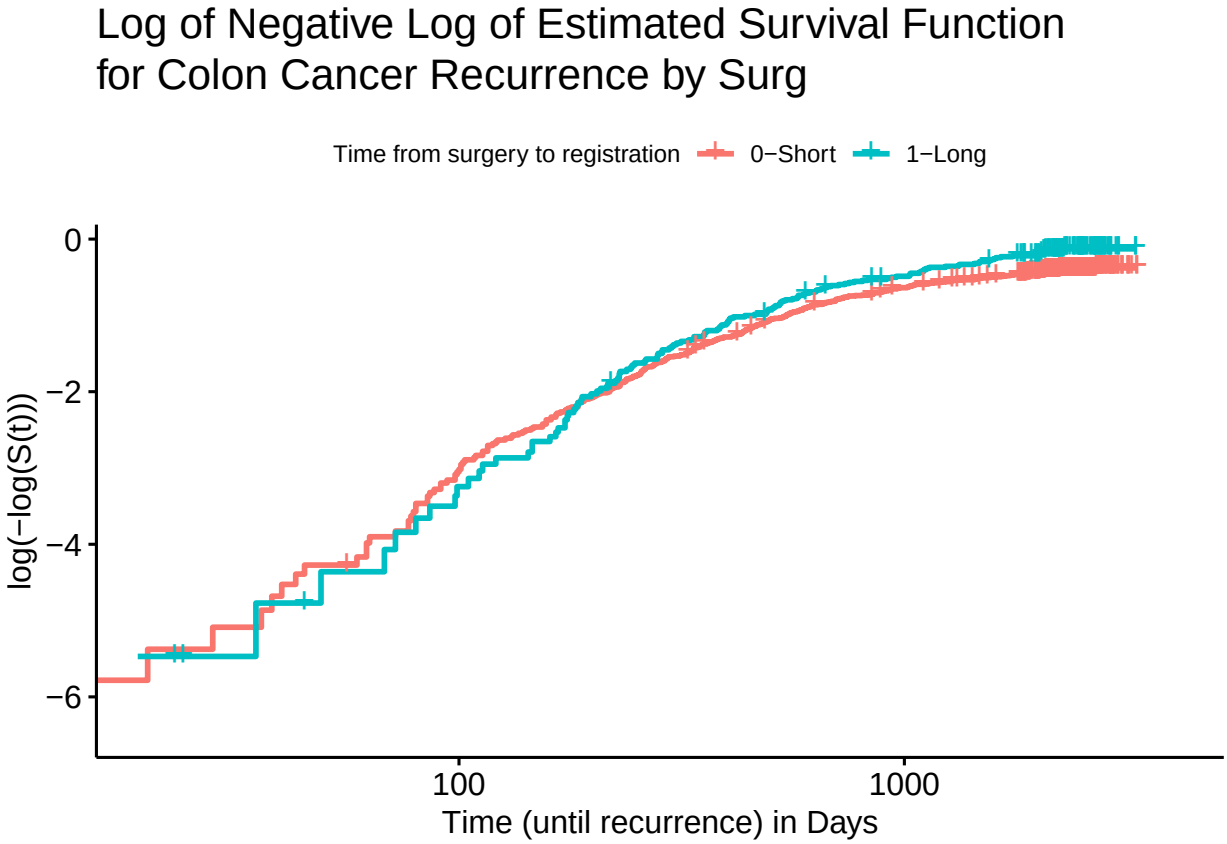
```
r.extent.fit <- survfit(Surv(time, status) ~ extent, data = colon.recur)
cll.extent = ggsurvplot(r.extent.fit, conf.int = F, font.x.size = 12, font.y.size = 12, font.legend.size
           fun = "cloglog",
           xlim = c(20, 4000),
           xlab = "Time (until recurrence) in Days",
           legend.lab = c("1-Submucosa","2-Muscle","3-Serosa", "4-Contiguous structures"),
           legend.title = "Extent of local spread",
           title = "Log of Negative Log of Estimated Survival Function \nfor Colon Cancer Recurrence by
cll.extent
```



According to the plot, the curves in the C-log-log plot are crossing with each other. The assumption of proportional hazard of extent is violated as we fail to see four parallel lines against log time.

We proceed with the C-log-log plot for the covariate surg.

```
r.surg.fit <- survfit(Surv(time, status) ~ surg, data = colon.recur)
cll.surg = ggsurvplot(r.surg.fit, conf.int = F, font.x.size = 12, font.y.size = 12, font.legend.size = 
           fun = "cloglog",
           xlim = c(20, 4000),
           xlab = "Time (until recurrence) in Days",
           legend.lab = c("0-Short","1-Long"),
           legend.title = "Time from surgery to registration",
           title = "Log of Negative Log of Estimated Survival Function \nfor Colon Cancer Recurrence by
cll.surg
```

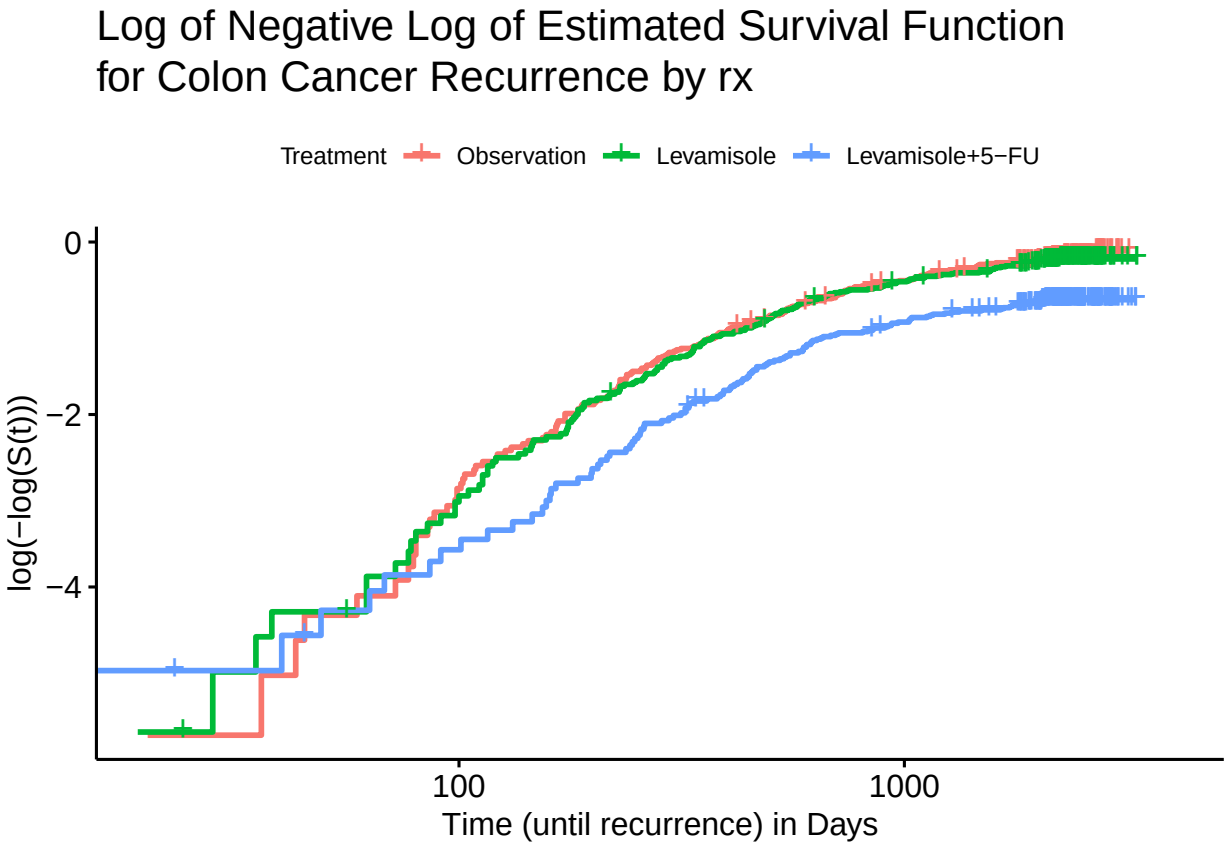


According to the C-log-log plot for surg, the lines of two groups coincide, which indicates that the cox proportional hazards assumption holds for surg.

We continue to plot the C-log-log plot for the covariate rx.

```
r.rx.fit <- survfit(Surv(time, status) ~ rx, data = colon.recur)
cll.rx <- ggsurvplot(r.rx.fit, conf.int = F, font.x.size = 12, font.y.size = 12, font.legend.size = 9,
                          fun = "cloglog",
                          xlim = c(20, 4000),
                          xlab = "Time (until recurrence) in Days",
                          title = "Log of Negative Log of Estimated Survival Function \nfor Colon Cancer
                          legend.title = "Treatment",
                          legend.labs = c("Observation", "Levamisole","Levamisole+5-FU"))

cll.rx
```



According to the plot, the curves for observation group and levamisole group coincide, while the curve for levamisole + 5-FU is parallel to the other two curves. The proportional hazard assumption holds for covariate rx.

```
splots <- list()
splots[[1]] <- cll.differ
splots[[2]] <- cll.extent
splots[[3]] <- cll.surg
splots[[4]] <- cll.rx


cloglog_plot = arrange_ggsurvplots(splots, print = FALSE, ncol = 2, nrow = 2)
ggsave(cloglog_plot,file = "./plot/r.C-log-log-plots.pdf",width = 12,height = 15)
ggsave(cloglog_plot,file = "./plot/r.C-log-log-plots.png",width = 12,height = 15)
```

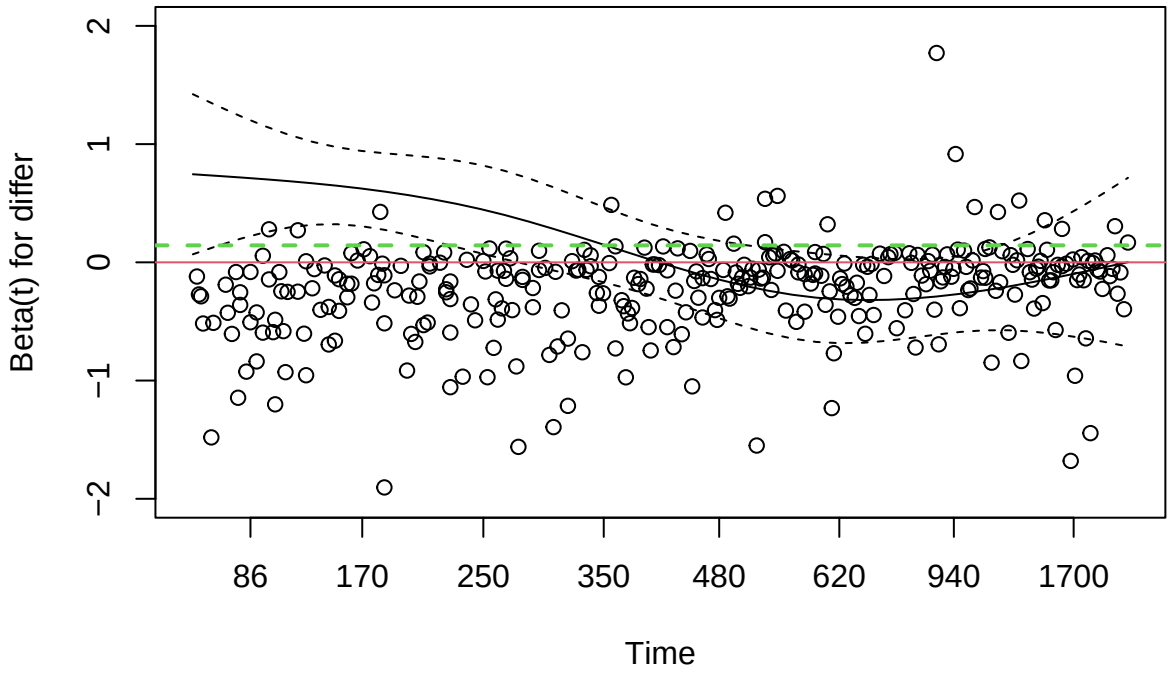
Schoenfeld residuals

```
r.residual.fit = coxph(Surv(time, status) ~ differ + rx + extent + surg + nodes, data = colon.recur)
r.residual = cox.zph(coxph(Surv(time, status) ~ differ + rx + extent + surg + nodes, data = colon.recur
r.residual
```

```
##         chisq df       p
## differ 13.525  1 0.00024
## rx      0.454  2 0.79696
## extent  0.140  1 0.70802
## surg    1.827  1 0.17644
```

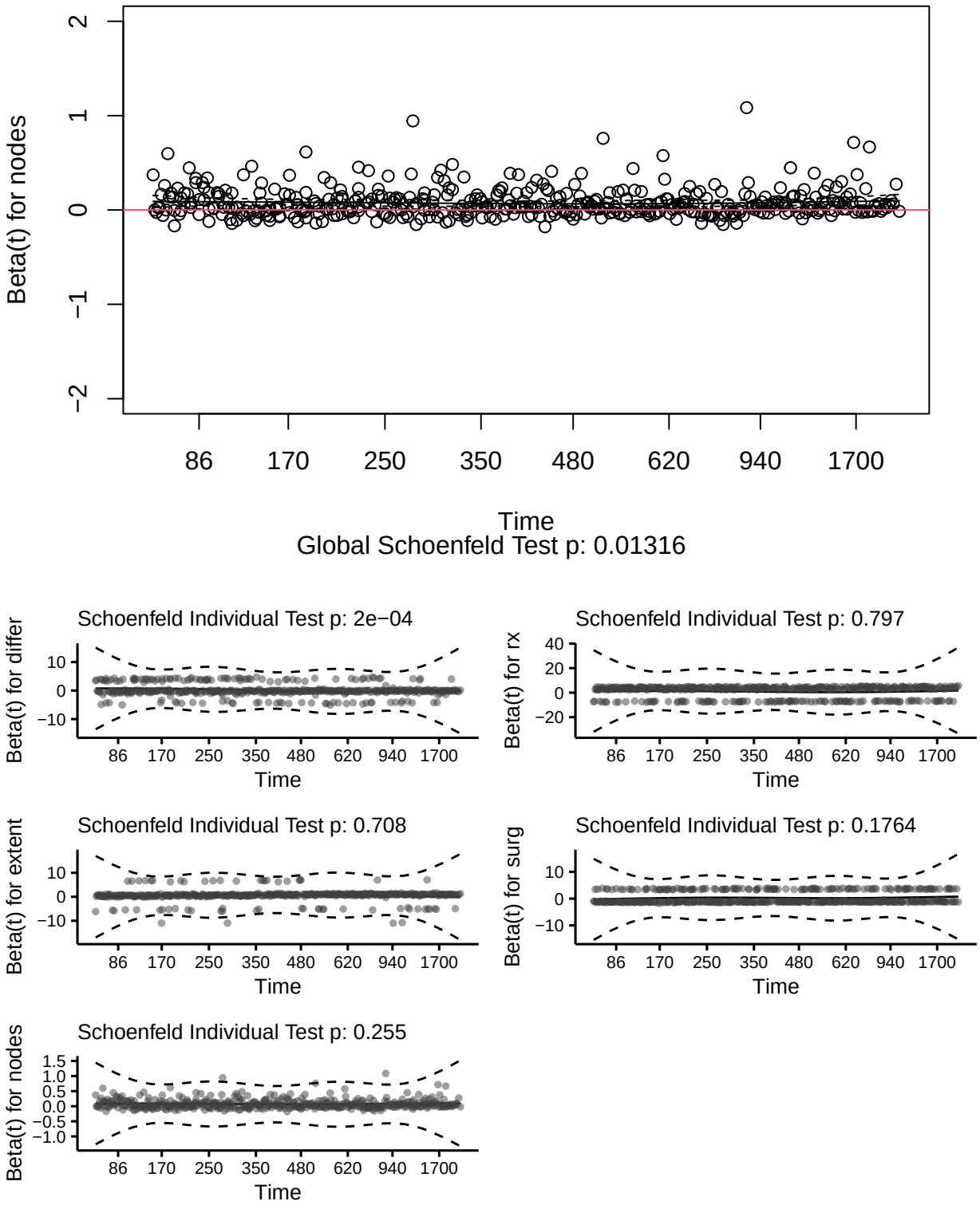
```
## nodes   1.295  1 0.25504
## GLOBAL 16.112  6 0.01316
```

```
# residual plot for covariate that violate ph assumption
plot(r.residual[1], ylim = c(-2, 2))
abline(a = 0, b = 0, col = 2)
abline(h = r.residual.fit$coef[1], col = 3, lwd = 2, lty = 2)
```




```
plot(r.residual[5], ylim = c(-2, 2))
abline(a = 0, b = 0, col = 2)
abline(h = r.residual.fit$coef[8], col = 3, lwd = 2, lty = 2)


residual_plot = ggcoxzph(r.residual, font.main = 10, font.x = 10, font.y = 10, font.tickslab = 8,
         point.alpha = 0.5, point.col = "grey25")
residual_plot
```




```
ggsave("./plot/r.schoenfeld residual_plots.pdf", arrangeGrob(grobs = residual_plot))
ggsave("./plot/r.schoenfeld residual_plots.png", arrangeGrob(grobs = residual_plot))
```

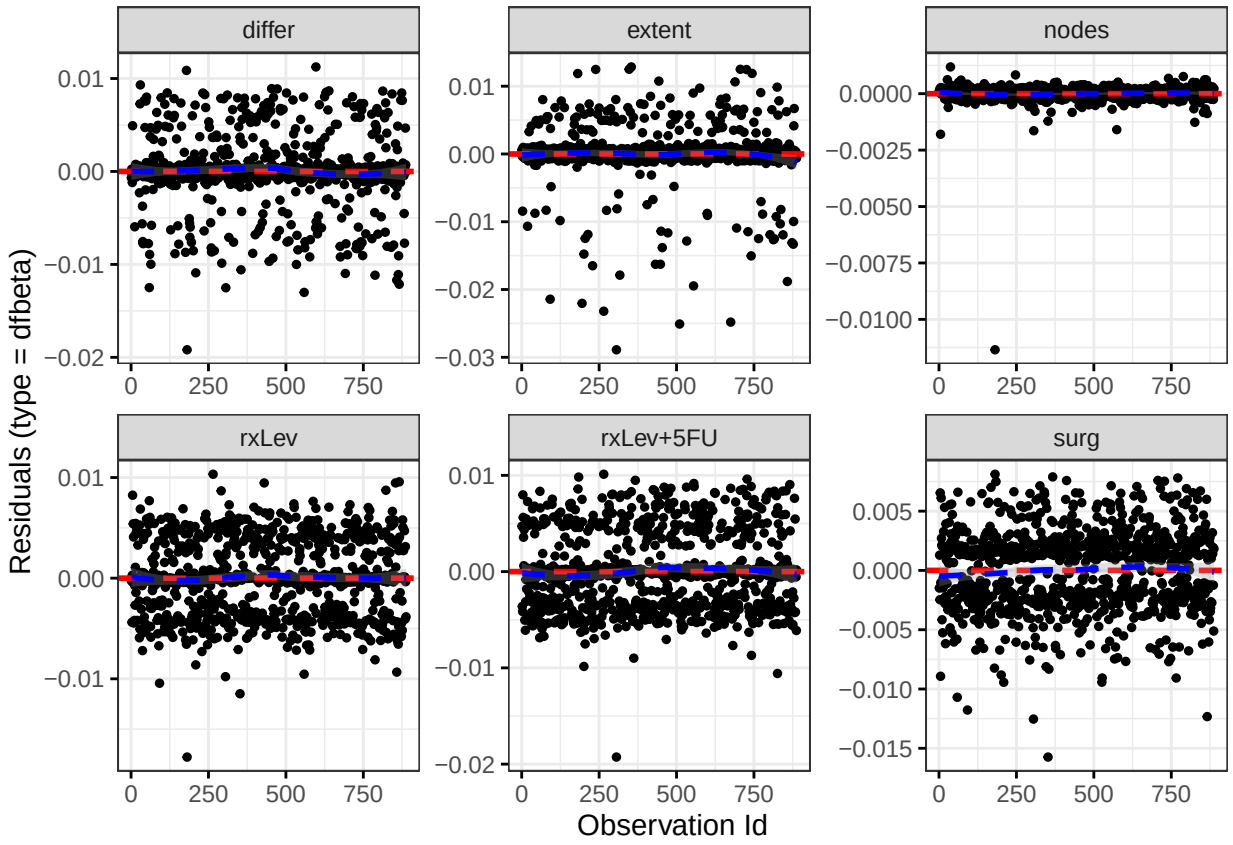
differ violates ph assumption

From the output above, the tests for covariates “differ” is statistically significant. Therefore, we assume there

is violation of proportional hazard assumption on "differ", which requires correction of non-proportionality.

4.2.2 Test influential observations

```
# check outliers in term of dfbeta
r.outlier_dfbeta <- ggcoxdiagnostics(coxph(Surv(time, status) ~ differ + extent + surg + nodes + rx, da
                                    type = "dfbeta", linear.predictions = FALSE, ggtheme = theme_bw())
r.outlier_dfbeta
```




```
ggsave(r.outlier_dfbeta, file = "./plot/r.outlier_dfbeta.pdf")
ggsave(r.outlier_dfbeta, file = "./plot/r.outlier_dfbeta.png")
```

In terms of dfbeta, the estimated changes in the regression coefficients upon deleting each observation, comparing the magnitudes of the largest dfbeta values to the regression coefficients suggests that none of the observations is terribly influential individually, even though some of the dfbeta values for extent are large compared with the others.

4.3 Modification for Violation of PH Assumption

```
# add interaction of covariate with function of time
r.model.inter = coxph(Surv(time, status) ~ rx + extent + surg + differ + tt(differ) + nodes, data = col
anova(r.model.inter)
```

```
## 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         -4997.3 -4240.3845  2   1.000000
## extent     -4978.0    38.5619  1  5.304e-10 ***
## surg       -4975.0     6.1549  1   0.013105 *
## differ     -4971.3     7.2540  1   0.007074 **
## tt(differ) -4335.5  1271.6363  1  < 2.2e-16 ***
## nodes      -2815.3  3040.4929  1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(r.model.inter)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ rx + extent + surg + differ +
##     tt(differ) + nodes, data = colon.recur, tt = function(x,
##     t, ...) x * log(t))
##
##   n= 888, number of events= 446
##
##                 coef exp(coef)  se(coef)      z Pr(>|z|)
## rxLev      -0.068013  0.934248  0.110681 -0.614 0.538887
## rxLev+5FU  -0.508518  0.601386  0.121490 -4.186 2.84e-05 ***
## extent      0.501526  1.651240  0.116837  4.293 1.77e-05 ***
## surg        0.240277  1.271602  0.103884  2.313 0.020726 *
## differ      2.228035  9.281609  0.583345  3.819 0.000134 ***
## tt(differ) -0.353111  0.702499  0.097562 -3.619 0.000295 ***
## nodes       0.075754  1.078697  0.009329  8.121 4.64e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##            exp(coef) exp(-coef) lower .95 upper .95
## rxLev         0.9342     1.0704    0.7521    1.1606
## rxLev+5FU     0.6014     1.6628    0.4740    0.7631
## extent        1.6512     0.6056    1.3133    2.0762
## surg          1.2716     0.7864    1.0373    1.5588
## differ        9.2816     0.1077    2.9586   29.1184
## tt(differ)    0.7025     1.4235    0.5802    0.8505
## nodes         1.0787     0.9270    1.0592    1.0986
##
## Concordance= 0.661  (se = 0.013 )
## Likelihood ratio test= 123.7  on 7 df,   p=<2e-16
## Wald test            = 141.9  on 7 df,   p=<2e-16
## Score (logrank) test = 144.7  on 7 df,   p=<2e-16
```

Since we have observed the violation of PH assumption for differ from Schoenfeld residuals, we modified the model by adding an interaction of covariate differ with transformation of time.

4.4 Final model

With correction terms, our final model is given by: Surv(time, status) ~ rx + extent + surg + differ + tt(differ) + nodes.

```
r.model.final = coxph(Surv(time, status) ~ rx + extent + surg + differ + tt(differ) + nodes, data = colo
summary(r.model.final)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ rx + extent + surg + differ +
##     tt(differ) + nodes, data = colon.recur, tt = function(x,
##     t, ...) x * log(t))
##
##   n= 888, number of events= 446
##
##                  coef exp(coef)  se(coef)      z Pr(>|z|)
## rxLev      -0.068013  0.934248  0.110681 -0.614 0.538887
## rxLev+5FU  -0.508518  0.601386  0.121490 -4.186 2.84e-05 ***
## extent      0.501526  1.651240  0.116837  4.293 1.77e-05 ***
## surg        0.240277  1.271602  0.103884  2.313 0.020726 *
## differ      2.228035  9.281609  0.583345  3.819 0.000134 ***
## tt(differ) -0.353111  0.702499  0.097562 -3.619 0.000295 ***
## nodes       0.075754  1.078697  0.009329  8.121 4.64e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##            exp(coef) exp(-coef) lower .95 upper .95
## rxLev         0.9342     1.0704    0.7521    1.1606
## rxLev+5FU     0.6014     1.6628    0.4740    0.7631
## extent        1.6512     0.6056    1.3133    2.0762
## surg          1.2716     0.7864    1.0373    1.5588
## differ        9.2816     0.1077    2.9586   29.1184
## tt(differ)    0.7025     1.4235    0.5802    0.8505
## nodes         1.0787     0.9270    1.0592    1.0986
##
## Concordance= 0.661  (se = 0.013 )
## Likelihood ratio test= 123.7  on 7 df,   p=<2e-16
## Wald test            = 141.9  on 7 df,   p=<2e-16
## Score (logrank) test = 144.7  on 7 df,   p=<2e-16
```

5. Conclusion

Recall that our final model is given by: Surv(time, status) ~ rx + extent + surg + differ + tt(differ) + nodes.

Since the p-value for the coefficient of rxLev group is 0.539 which is larger than 0.05, there is not significant evidence to indicate a difference of tiem of recurrence between the group treated with rxLev and the observation group. Additionally, we are 95% confident that the true risk ratio of recurrence is between [0.752, 1.161]. Therefore, we conclude that the treatment rxLev is not effective on decreasing the risk of recurrence.

On the other hand, our result indicates that the hazard ratio for the group treated with rxLev + 5FU is 0.601, corresponding to a 40% decrease in risk of recurrence compared to the observation group. Since the p-value is 0.0000284 which is smaller than 0.05, there is a significant difference between the group treated with rxLev + 5FU and the observation group. Additionally, corresponding 95% CI is between [0.474, 0.763]. Therefore, we conclude that the treatment rxLev + 5FU is effective on decreasing the risk of recurrence.

Therefore, we conclude that the treatment levamisole + 5-FU is effective as an adjuvant chemotherapy on decreasing the risk of recurrence.