Survival analysis with Malignant Melanoma

Liela Meng

Department of Biostatistics, University of California, Davis, ONE SHIELDS AVENUE, Davis, 95616, California, United States. Tel.: 5309336513

Email: mzcmeng@ucdavis.edu (Liela Meng)

Abstract

Aim of the study: Non-parametric methods and the Cox PH model were applied to investigate whether female melanoma patients have a lower risk of death than males.

Study population: Patient had his or her melanoma tumor removed at the Department of Plastic Surgery, University Hospital of Odense, Denmark during the period 1962 to 1977 were included.

Main variables: The main variables include demographic, clinical, and pathological characteristics, including survival time since the operation, patient status at the end of the study, sex, age, year of operation, tumor thickness in mm, and indicator of ulceration. In case of death, the cause of whether due to melanoma or not were recorded.

Results: Though the two-sample hypothesis test showed a significant difference between the two genders, the CoxPH model, after adjusted for age, tumor thickness, and ulceration of the tumor, indicated a non-significant survival difference between males and females patients.

Keywords: Survival analysis, Malignant Melanoma, Cox PH model, Prognostic factors

Survival analysis with Malignant Melanoma

Introduction

Malignant melanoma is the most dangerous type of skin cancer. There are several known prognostic factors have been discovered for the survival of melanoma patients. Thicker tumor, deeper invasion, poorer prognosis, localization of the tumor on the mucous membranes, ulceration, presence of epithelioid cells are all bad prognostic signs. Females tend to have a better prognosis than males and the prognosis aggravates with age(Drzewiecki (1982)).

Objective

The current analysis aimed at investigating after tumor removal surgery, whether female malignant melanoma patients have a lower risk of death from melanoma than male patients.

Background information of the study

Each patient had his or her tumor removed at the Department of Plastic Surgery, University Hospital of Odense, Denmark during the period 1962 to 1977. The surgery consisted of the complete removal of the tumor together with about 2.5cm of the surrounding skin. All patients were followed until the end of 1977.

Characteristics of Data

The total number of patients included in this analysis is 205. Among the measurements taken were the thickness of the tumor and whether it was ulcerated or not. Besides, the sex and age of patients, the operation year, and patients' status at the end of the study were also recorded.

The following variables were provided in the data set: status (died from melanoma, alive, or died from other reasons), sex, age, thickness (in mm), ulcer (present or not), year (of operation), time to event. Overall follow-up time was considered from the date of operation to the date of death (event of interest) or censoring (alive at the end of 1997 or died from other reasons).

To sum up, the time origin is the date of tumor removal surgery, and the event of interest is death from melanoma and recorded in years. This is a generalized type I right censoring with random censoring data set.

Statistical Analysis

Descriptive analysis

Descriptive statistical analysis was carried out to explore the patient and disease characteristics using mean$\pm$ SD and median for continuous variables, frequency for categorical variables.

Non-parametric estimation

To initially analyze the data, the Kaplan-Meier (KM) estimator was used to quantify the mean survival time based on gender. Smoothed hazard rates were plotted to graphically detect survival time differences for both genders. Hypothesis testings were also conducted to quantify the difference in survival time between the two genders.

Assumption of KM method: non-informative censoring. In other words, censoring time is independent of survival time.

The probability of being alive at time t_i, $s(t_i)$is calculated by

$\prod_{t_i \leqslant t}(1 - \frac{d_i}{Y_i})$, where $t_0 = 0,\ S(0) = 1$. Y_i is the number of patients alive just before t_i and d_i is the number of event at t_i.

Cox PH model

Then, the proportional Cox regression model was used to analyze the survival time to adjust for other covariates. Using forward procedure and AIC, local test p-value as criteria, a CoxPH model was built. Local hypothesis testings were used to check whether any interaction terms should be included in the model. Later, model diagnose were conduct to decide the proper form of each covariate include in the model, check overall fitness of the final model, assess proportional hazard assumption, and identify outliers. The stratified CoxPH model was contrasted to address the violation of the PH assumption problem. Lastly, a likelihood ratio test was used to assess whether each stratum shares the same regression coefficient.

A significant level of 0.05 was used. All statistical analyses were conducted using SAS University Edition.

Assumptions of Cox PH model: 1)Independent identically distributed observations; 2) non-informative censoring; 3) constant hazard ratio over time; 4) effect of covariates is additive and linear on a log-hazard scale.

Additional assumptions for stratified Cox PH model: no interaction. In other words, each stratum shares the same regression coefficients.

Results

Characteristics of Prognostic Variables

A total of 205 patients with Malignant Melanoma were included in the current analysis. The $mean \pm SD$ for age at the time of operation was 52.46 ± 16.67 years. In addition, the $mean \pm SD$ for tumor thickness was 2.92 ± 2.96 mm, and 44% of them had ulcers in the tumor. In the study, 57 (27.8%) patients died of melanoma until the end of 1977 (see Table 1). Based on the life-table method, the probability of surviving beyond one year was 97%. The year of operation was categorized into a three-level categorical variable with breakpoints year 1967 and the year 1972.

Table 1 shows the characteristics of prognostic variables for female and male patients. For both genders, age, thickness, and operation year were similar. But ulcer of the tumor was significantly different for the two genders. Thus, an interaction between sex and ulcer would be considered later.

Figure C2 shows overall survival probability, 95% confidence band, and point-wise confidence interval using the kaplan-Meier (KM) method. Figure C1 is the smoothed hazard rate plot. The hazard rate increase at the very beginning to year 3, then gradually decreases till year 8, after that, the hazard rate continues to increase. Censored observations are represented by vertical ticks on the graph. Since the observation with the longest follow-up is censored, the survival function will not reach 0. Due to heavy censoring after 4 years follow-up time, the median survival time was undefined and the mean survival time based on the KM method was 7.4 years.

Stratified by Gender

Of the 205 patients, 126 (61%) of them were female, and 79 (39%) were male. In female patients, 28 (22%) were dead due to melanoma with a mean survival time of 7.8 years. In male patients, 29 (37%) were dead due to melanoma with a mean survival time of 5.8 years.

Since the primary goal of this study was to investigate the survival experience of different gender, stratified KM survival estimate curves were presented. From Figure 1, the KM curve for females is constantly above males, indicating overall females have a better survival experience than males. This difference is reinforced by the three two-sample tests based on gender (Table C1). Stratified tests of gender after adjusted for ulcer were also conducted and all of those tests indicate survival difference between the two genders.

From Figure C3, the smoothed hazard function indeed is higher for males compared to females. A female hazard rate increase at the beginning then staying mostly constant from 3 to 8 years follow-up time, and then increase again. In contrast, the male hazard rate remains constantly high from the beginning to 3 years follow-up time, then decreasing till 7 years follow-up time. The sudden upticks at the end of follow-up time were not reliable since the number of subjects at risk was too small.

CoxPH Model

Model Building

Using the forward method by adding one variable at each time, choosing minimal AIC and P value<0.05 as inclusion criteria, variables included in the model are: sex (always included), thickness, and ulcer. Besides, since age has been shown as a significant prognostic factor in the former article (Drzewiecki (1982)), age was also included in the model (final model AIC =532.78). No interaction terms showed significance in local hypothesis tests, thus the model does not include any interaction terms.

Table C2 shows after adjusted for ulcer, thickness, and age, on average, a patient who is female approximately has a 45% lower risk of dying from melanoma compared to male patients, but this decrease is not significant (HR=0.65, 95%CI:[0.38, 1.10], P-value=0.11).

Functional Forms of continuous covariates

Martingale residuals is a measure of excess observed events and smooth of a scatter plot of the martingale residuals approximate the correct functional form of a covariate. Based on martingale residuals, variable thickness of tumor was suggested to be categorized into a three levels categorical variable with cut points 4.84 mm and 7.41 mm (Figure C4). However, since numbers at risk for levels 2 and 3 were too small, at last, only cut point 4.84 mm was used. Figure C5 shows no need to categorize age.

Assessing the proportional hazards assumption

Figure C10 shows, for the thickness level of the tumor, two curves are not parallel. Schoenfeld residuals plot Figure C6 , having a decreasing line, also suggests the same conclusion. Thus, the PH assumption does not hold for thickness, and stratified CoxPH models are required.

Though Figure C7 residual plot of age shows a slightly positively trending at the end, but the smooths appear mostly near 0 before 4 years follow-up time (after which heavy censored). Thus, taking heavy censoring into consideration, I concluded the PH assumption holds for age.

Log-log survival plot of ulcer and sex Figures C8 and C9 seems to have parallel curves, thus no further change will make for those two variables.

Test for equal regression coefficient assumption had the result$LRT = 468.618 - 387.011 - 78.982 = 2.625,\ P_value = 0.45, df = 3$which means regression coefficients are the same in each stratum of thickness.

Figure C11 shows no outliers require special attention.

Final Model

For

$$h_k(t) = h_{0,k} \exp(\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3),\ k = 1\ or\ 2.$$

$h_{0,1}$ for stratum thickness of tumor less than 4.84mm. X_1, X_2, X_3 presents for sex, age and ulcer, respectively. The estimation of parameter betas is shown in Table C3.

Cox-Snell Residuals shows the overall fitness of the final model is not too bad (Figure C12).

Model Interpretation

After adjusting for age, sex, and thickness of the tumor, on average, a malignant melanoma patient without an ulcer on the tumor has a 74% lower risk of death from melanoma compared to patients who has an ulcer on the tumor, with 95% sure that the true value is lying between 52%-86%.

As for gender, though the two-sample test of gender with or without adjusted for ulcer showed a significant difference between the two genders. Cox PH model, after adjusting for age, ulcer, and thickness, indicated no significant effect of sex on survival experience for melanoma patients.

Discussion

Though the stratified Cox PH model's no interaction assumption holds thus and gave us sufficient evidence to show ulcer was a significant prognostic factor, this study still suffers from several drawbacks:

There are other factors that might also affect melanoma patients' survival time and failure to obtain data regarding those variables that could bias the result of the Cox PH model.

Due to heavy censoring, the number at risk after 6 years was very small. Thus, lots of problems are caused by this feature and could not be solved. For example, the Schoenfeld residual plot of age had an increase after 6 years, but we could not conclude age was a time-dependent variable based on this plot since data after 8 years were not reliable. Besides, The KM method is not ideal for heavy censoring data (censoring rate over 50%), thus the conclusion of the significant survival time difference between two genders also was not reliable.

Assumptions regarding the independence of censoring time and event time and independent observations were unable to know since the study design description did not provide such information and there is no existing test to test whether non-informative

assumption holds. But since melanoma is not a contagious disease and the reason for death has been recorded, it is safe to assume those two assumptions hold.

Conclusions

This study found one significant prognostic factor for malignant melanoma: whether there is an ulcer on the tumor or not. The primary hypothesis regarding gender affects the survival experience of melanoma patients could not be verified. Further study should gather more information especially for other known prognostic factors and prolong the follow-up time could reduce the bias caused by heavy censoring.

Risk factors	Levles	Overall N.(%)	Female	Male	P-value[a]
Age	Mean±SD,	52.46±16.67,	51.7±16.1,	53.9±17.6,	0.359
	median	54	54	55	
Thickness	Mean±SD,	2.92±2.96,	2.49±2.75,	3.61±3.16,	0.175
	median	1.94	1.62	2.58	
Ulcer	Absent	115 (56%)	79 (62.7%)	36 (45.6%)	0.016*
	Present	90 (44%)	47 (37.3%)	43 (54.4%)	
Year	1962-1967	43 (21.0%)	27 (21.4%)	16 (20.3%)	0.949
	1968-1972	129 (62.9%)	53 (42.1%)	35 (44.3%)	
	1972-1977	33 (16.1%)	46 (36.5%)	28 (35.4%)	

[a]: P-values for quantitative variables were obtained by t test for gender. P-values for categorical variables were obtained by Chi-squared test for equal proportions for gender.

*: Indicate statistically significant at level 0.05

Table 1

Prognostic variables of the 205 patients had melanoma tumor removing surgery

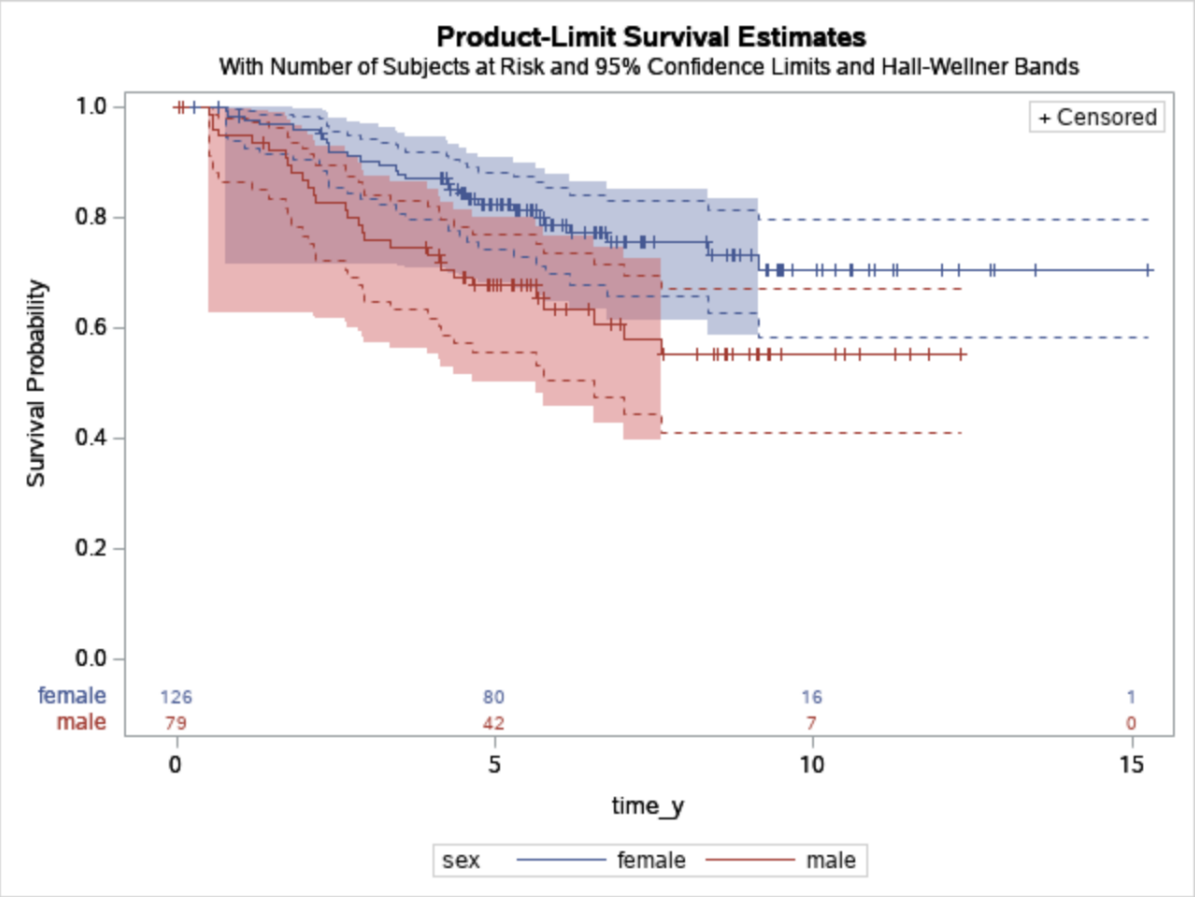

Figure 1

Stratified KM survival estimates based on sex

Appendix A

Appendix

Appendix B

Appendix - SAS code

```
LIBNAME project1 "/folders/myfolders/BST222/project1";

        proc import datafile="/folders/myfolders/BST222/project1/melanoma.csv"
out=project1.data0 dbms=csv;
getnames=yes;
run;

        data project1.data(drop= var1 );
set project1.data0;
if status=1 then censored=0;
else censored=1;
if year<=1967 then do; yearcate=1; yearcate1=0; yearcate2=0 ;end;
else if 1967<year<=1971 then do; yearcate=2 ;yearcate1=1; yearcate2=0 ;end;
else if year>1971 then do; yearcate=3; yearcate1=0; yearcate2=1 ;end;
id=var1;
run;

        Proc format;
Value status
1 = 'died from melanoma'
2 = 'still alive'
3 = 'died not from melanoma';
Value ulcer
0 = 'absent'
1 = 'present';
Value sex
0 = 'female'
1 = 'male';
Value censored
0 = 'not censored'
```

```
1 = ’censored’;
value thick
1 = ’thickness<=4.84mm’
2 = ’thickness > 4.84mm’;
Run;
* discretize thickness;
data project1.datanew(drop=time );
set project1.data;
time_y=time/365;
if thickness<=4.84 then thick=1;
else if 4.84<thickness then thick=2;
run;

        ************************
Table 1
************************;
proc freq data=project1.datanew;
table sex;
format sex sex. censored censored.;
run;
proc sort data=project1.datanew;
by sex;
run;
proc freq data=project1.datanew;
tables censored;
by sex;
format sex sex. censored censored.;
run;
* equality test;
proc univariate data=project1.datanew;
```

```
class sex;
var age thickness;
format sex sex.;
run;
proc ttest data=project1.datanew;
class sex;
var age thickness;
format sex sex.;
run;
proc freq data=project1.datanew;
tables (ulcer yearcate)*sex/ chisq norow;*Trend cmh;
format sex sex. ulcer ulcer.;
run;
**********
Non parametric
**********;
*life table;
proc LIFETEST data=project1.datanew method=lt atrisk intervals=(0 to 20 by 1);
time time_y*censored(1);
format sex sex. censored censored.;
run;
* overall;
proc LIFETEST data=project1.datanew plots=survival(cl cb=hw strata=overlay atrisk) ;
time time_y*censored(1);
format sex sex. censored censored.;
run;
proc LIFETEST data=project1.datanew plots=hazard notable;
time time_y*censored(1);
format sex sex. censored censored.;
```

```
run;
* by sex;
proc LIFETEST data=project1.datanew plots=survival(cl cb=hw strata=overlay atrisk) ;
time time_y*censored(1);
strata sex;
format sex sex. censored censored.;
run;
proc LIFETEST data=project1.datanew plots=hazard notable;
time time_y*censored(1);
strata sex;
format sex sex. censored censored.;
run;
************************
two sample test
************************;
* two sample test - fleming (1,0);
proc lifetest data=project1.datanew notable;
time time_y*censored(1);
strata sex/test=(wilcoxon fleming(1,0) logrank );
format sex sex.;
run;
* statified test - ulcer;
proc lifetest data=project1.datanew notable;
time time_y*censored(1);
strata sex/group=ulcer test=(wilcoxon fleming(1,0) logrank);
format sex sex. ulcer ulcer.;
run;
************************
Local test- interaction
```

```
**********;
*sex;
proc phreg data=project1.datanew;
class sex;
model time_y*censored(1)= sex age age*sex;
contrast "beta3=0" sex*age 1 /test(wald lr score);
format sex sex. ;
run;
proc phreg data=project1.datanew;
class sex;
model time_y*censored(1)= sex thickness sex*thickness;
contrast "beta3=0" sex*thickness 1 /test(wald lr score);
format sex sex. ;
run;
proc phreg data=project1.datanew;
class sex ulcer;
model time_y*censored(1)= sex ulcer sex*ulcer;
contrast "beta3=0" sex*ulcer 1 /test(wald lr score);
format sex sex. ulcer ulcer.;
run;
proc phreg data=project1.datanew;
class sex yearcate;
model time_y*censored(1)= sex yearcate sex*yearcate;
contrast "beta3=beta4=0" sex*yearcate 1 0,
sex*yearcate 0 1/test(wald lr score);
format sex sex. ;
run;
*thickness;
proc phreg data=project1.datanew;
```

```
model time_y*censored(1)= thickness age age*thickness;
contrast "beta3=0" thickness*age 1 /test(wald lr score);
run;
proc phreg data=project1.datanew;
class ulcer;
model time_y*censored(1)= thickness ulcer thickness*ulcer;
contrast "beta3=0" thickness*ulcer 1 /test(wald lr score);
format ulcer ulcer.;
run;
proc phreg data=project1.datanew;
class yearcate;
model time_y*censored(1)= thickness yearcate thickness*yearcate;
contrast "beta3=beta4=0" thickness*yearcate 1 0,
thickness*yearcate 0 1/test(wald lr score);
run;
*ulcer;
proc phreg data=project1.datanew;
class ulcer;
model time_y*censored(1)= ulcer age ulcer*age;
contrast "beta3=0" ulcer*age 1 /test(wald lr score);
format ulcer ulcer.;
run;
proc phreg data=project1.datanew;
class ulcer yearcate;
model time_y*censored(1)= ulcer yearcate ulcer*yearcate;
contrast "beta3=beta4=0" ulcer*yearcate 1 0,
ulcer*yearcate 0 1/test(wald lr score);
format ulcer ulcer.;
run;
```

```
*age;
proc phreg data=project1.datanew;
class yearcate;
model time_y*censored(1)= age yearcate age*yearcate;
contrast "beta3=beta4=0" age*yearcate 1 0,
age*yearcate 0 1/test(wald lr score);
format ulcer ulcer.;
run;
**********
Model selection
**********;
* sex +;
proc phreg data=project1.datanew plots(overlay)=(survival);
model time_y*censored(1)= sex age;
age: test age = 0;
format sex sex. ;
ods output fitstatistics = ageaic;
ods output teststmts = agestat;
run;
proc phreg data=project1.datanew plots(overlay)=(survival);
model time_y*censored(1)= sex thickness;
thickness:test thickness=0;
format sex sex. ;
ods output fitstatistics = ageaic;
ods output teststmts = agestat;
run;
proc phreg data=project1.datanew plots(overlay)=(survival);
model time_y*censored(1)= sex ulcer;
ulcer: test ulcer = 0;
```

```
format sex sex. ulcer ulcer.;
ods output fitstatistics = ulceraic;
ods output teststmts = ulcerstat;
run;
proc phreg data=project1.datanew plots(overlay)=(survival);
class yearcate;
model time_y*censored(1)= sex yearcate;
contrast "beta2=beta3=0" yearcate 1 0,
yearcate 0 1/test(wald lr score);
format sex sex.;
ods output fitstatistics = yearcateaic;
run;
* +ulcer;
proc phreg data=project1.datanew plots(overlay)=(survival);
model time_y*censored(1)= sex ulcer thickness;
thickness: test thickness = 0;
format sex sex. ulcer ulcer.;
run;
proc phreg data=project1.datanew plots(overlay)=(survival);
model time_y*censored(1)= sex ulcer age;
age: test age = 0;
format sex sex. ulcer ulcer.;
run;
proc phreg data=project1.datanew plots(overlay)=(survival);
class yearcate;
model time_y*censored(1)= sex ulcer yearcate;
contrast "beta3=beta4=0" yearcate 1 0,
yearcate 0 1/test(wald lr score);
format sex sex. ulcer ulcer.;
```

```
run;
* +thickness;
proc phreg data=project1.datanew plots(overlay)=(survival);
model time_y*censored(1)= sex ulcer thickness age;
age: test age = 0;
format sex sex. ulcer ulcer.;
run;
proc phreg data=project1.datanew plots(overlay)=(survival);
class yearcate;
model time_y*censored(1)= sex ulcer thickness yearcate;
contrast "beta3=beta4=0" yearcate 1 0,
yearcate 0 1/test(wald lr score);
format sex sex. ulcer ulcer.;
run;
* final model (sex, ulcer, thickness, age);
proc phreg data=project1.datanew plots(overlay)=(survival);
class sex ulcer;
model time_y*censored(1)= sex age thickness ulcer/rl;
format sex sex. ulcer ulcer.;
run;
**********
Diagnostics
**********;
*martingale;
*thickness;
proc phreg data=project1.datanew;
model time_y*censored(1)= sex age ulcer ;
output out=plot2_1 RESMART = mgale ;
format sex sex. ulcer ulcer.;
```

```
run;
proc loess data=plot2_1;
model mgale = thickness / direct;
run;
* age;
proc phreg data=project1.datanew;
model time_y*censored(1)= sex ulcer thickness;
output out=plot2_1 RESMART = mgale ;
format sex sex. ulcer ulcer.;
run;
proc loess data=plot2_1;
model mgale = age / direct;
run;
* final model;
proc phreg data=project1.datanew plots=survival;
class thick ulcer sex;
model time_y*censored(1)= sex age ulcer thick/rl;
format sex sex. ulcer ulcer. thick thick.;
run;
**********
PH assumption
**********;
********** thick;
* log(H) vs time;
data new ;
set project1.datanew ;
cons = 1;
run;
proc phreg data = new;
```

```
class thick/param=ref;
model time_y*censored(1) =cons/rl ;
strata thick;
output out = base logsurv = ls /method = ch;
run;
data base;
set base ;
logH = log (-ls);
if thick= 1 then logH1 = logH;
else if thick= 2 then logH2 = logH;
proc sort;by thick time_y ;
proc print; var thick time_y logH logH1 logH2 ;
run;
proc sgplot data =base;
where logH ne .;
series x=time_y y=logH /group=thick ;
format thick thick.;
run;
* shoenfeld;
proc phreg data = project1.datanew;
class sex ulcer thick/param=ref;
model time_y*censored(1) = thick age sex ulcer;
output out = schoen ressch= schthick ;
format sex sex. ulcer ulcer. thick thick.;
run;
proc loess data=schoen;
model schthick=time_y/smooth=(0.2 0.4 0.6 0.8);
run;
********** age;
```

```
* shoenfelf;
proc phreg data = project1.datanew;
class sex ulcer thick/param=ref;
model time_y*censored(1) = age sex ulcer;
output out = schoen ressch= schage ;
format sex sex. ulcer ulcer. thick thick.;
run;
proc loess data=schoen;
model schage=time_y/smooth=(0.2 0.4 0.6 0.8);
run;
********** ulcer;
data project1.datanew ;
set new ;
cons = 1;
run;
proc phreg data = new;
class ulcer/param=ref;
model time_y*censored(1) =cons/rl ;
strata ulcer;
output out = base logsurv = ls /method = ch;
run;
data base;
set base ;
logH = log (-ls);
if ulcer= 1 then logH1 = logH;
else if ulcer= 0 then logH2 = logH;
proc sort;by ulcer time_y ;
proc print; var ulcer time_y logH logH1 logH2 ;
run;
```

```
proc sgplot data =base;
where logH ne .;
series x=time_y y=logH /group=ulcer ;
format ulcer ulcer.;
run;
********** sex;
data new ;
set project1.datanew ;
cons = 1;
run;
proc phreg data = new;
class sex(ref="1")/param=ref;
model time_y*censored(1) =cons/rl ;
strata sex;
output out = base logsurv = ls /method = ch;
run;
data base;
set base ;
logH = log (-ls);
if sex= 1 then logH1 = logH;
else if sex= 0 then logH2 = logH;
proc sort;by sex time_y ;
proc print; var sex time_y logH logH1 logH2 ;
run;
proc sgplot data =base;
where logH ne .;
series x=time_y y=logH /group=sex ;
format sex sex.;
run;
```

```
********** Overall-Snell;
proc phreg data=project1.datanew;
class sex ulcer thick;
model time_y*censored(1)= age sex ulcer ;
strata thick;
format sex sex. ulcer ulcer. thick thick.;
output out=plot1_1 logsurv = logsurv1/method=ch;
run;
data plot1_1;
set plot1_1;
snell = -logsurv1;
cons = 1;
run;
proc phreg data = plot1_1;
model snell*censored(1) = cons;
output out = plot1_2 logsurv = logsurv2 /method = ch;
run;
data plot1_2;
set plot1_2;
cumhaz = - logsurv2;
run;
proc sort data = plot1_2;
by snell;
run;
proc sgplot data = plot1_2;
step y=cumhaz x=snell /MARKERFILLATTRS=(color="red");
lineparm x=0 y=0 slope=1; /** intercept, slope **/
label cumhaz = "Estimated Cumulative Hazard Rates";
label snell = "Residual";
```

```
run;
**********************************
stratified CoxPH assumption
**********************************;
proc phreg data=project1.datanew;
class sex ulcer thick;
model time_y*censored(1)= age sex ulcer /rl;
strata thick;
format sex sex. ulcer ulcer. thick thick.;
run;
proc phreg data=project1.datanew;
where thick=1;
class sex ulcer ;
model time_y*censored(1)= age sex ulcer;
format sex sex. ulcer ulcer. thick thick.;
run;
proc phreg data=project1.datanew;
where thick=2;
class sex ulcer ;
model time_y*censored(1)= age sex ulcer ;
format sex sex. ulcer ulcer. thick thick.;
run;
```

Appendix C

*

References

Bradburn, M. J. & Clark, T. G. (2003). Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods. *Multivariate data analysis*, 6. /Users/mengzichun/Zotero/storage/D6DIYP2Z/Bradburn and Clark - 2003 - Survival Analysis Part II Multivariate data analy.pdf.

Bradburn, M. J., Clark, T. G., Love, S. B., & Altman, D. G. (2003). Survival Analysis Part III: Multivariate data analysis – choosing a model and assessing its adequacy and fit. *British Journal of Cancer*, *89*(4), 605–611. /Users/mengzichun/Zotero/storage/J5LDYQJ4/Bradburn et al. - 2003 - Survival Analysis Part III Multivariate data anal.pdf. doi:10.1038/sj.bjc.6601120

Clark, T. G., Bradburn, M. J., Love, S. B., & Altman, D. G. (2003). Survival Analysis Part I: Basic concepts and first analyses. *British Journal of Cancer*, *89*(2), 232–238. /Users/mengzichun/Zotero/storage/2D7IAXWI/Clark et al. - 2003 - Survival Analysis Part I Basic concepts and first.pdf. doi:10.1038/sj.bjc.6601118

Drzewiecki, (K. (1982).

Hölmich, L. R., Klausen, S., Spaun, E., Schmidt, G., Gad, D., Svane, I. M., … Ibfelt, E. H. (2016). The Danish Melanoma Database. *Clinical Epidemiology*, *Volume 8*, 543–548. /Users/mengzichun/Zotero/storage/A3K6YIRM/Rosenkrantz Hölmich et al. - 2016 - The Danish Melanoma Database.pdf. doi:10.2147/CLEP.S99484

Survival with malignant melanoma: A regression analysis of prognostic factors. (1982), 6. /Users/mengzichun/Zotero/storage/V63Y73JC/1982 - Survival with malignant melanoma A regression ana.pdf.

Test	Chi-Square(unadjusted)	P value	Chi-Square(adjusted) [a]	P value
Wilcoxon	7.09	0.007*	27.17	<0.001*
Fleming (1,0)	7.41	0.008*	38.56	<0.001*
Log-rank	6.47	0.011*	26.37	<0.001*

*: Indicate statistically significant at level 0.05.

[a]: stratified test via adjusting for ulcer

Table C1

Test of Equality over Strata Gender

Paramter	Label	DF	Parameter Estimates	SE	Chi Square	P value	Hazard Ratio (95% CI)
Sex	Female	1	-0.433	0.267	2.620	0.106	0.649 [0.38, 1.1]
Age		1	0.012	0.008	2.162	0.142	1.012 [0.1, 1.03]
Thickness		1	0.109	0.038	8.336	0.004*	1.115 [1.04, 1.2]
Ulcer	Absent	1	-1.164	0.310	14.133	<0.001*	0.312 [0.17, 0.57]

*: Indicate statistically significant at level 0.05.

Table C2

Summary of CoxPH model with four covariates

Parameter	Label	Parameter Estimate	Chi-Square	P-value	Hazard Ratio (95% CI)
Sex	female	-0.505	3.47	0.006	0.603 [0.354, 1.027]
Age		0.01	1.58	0.21	1.01 [0.994, 1.027]
Ulcer	absent	-1.342	19.444	<0.001*	0.261 [0.144, 0.475]

*: indicated statistically significant at level 0.05

[a]: Stratified on thickness of tumor (break point 4.84mm)

Table C3

Summary of Stratified Cox PH model [a]

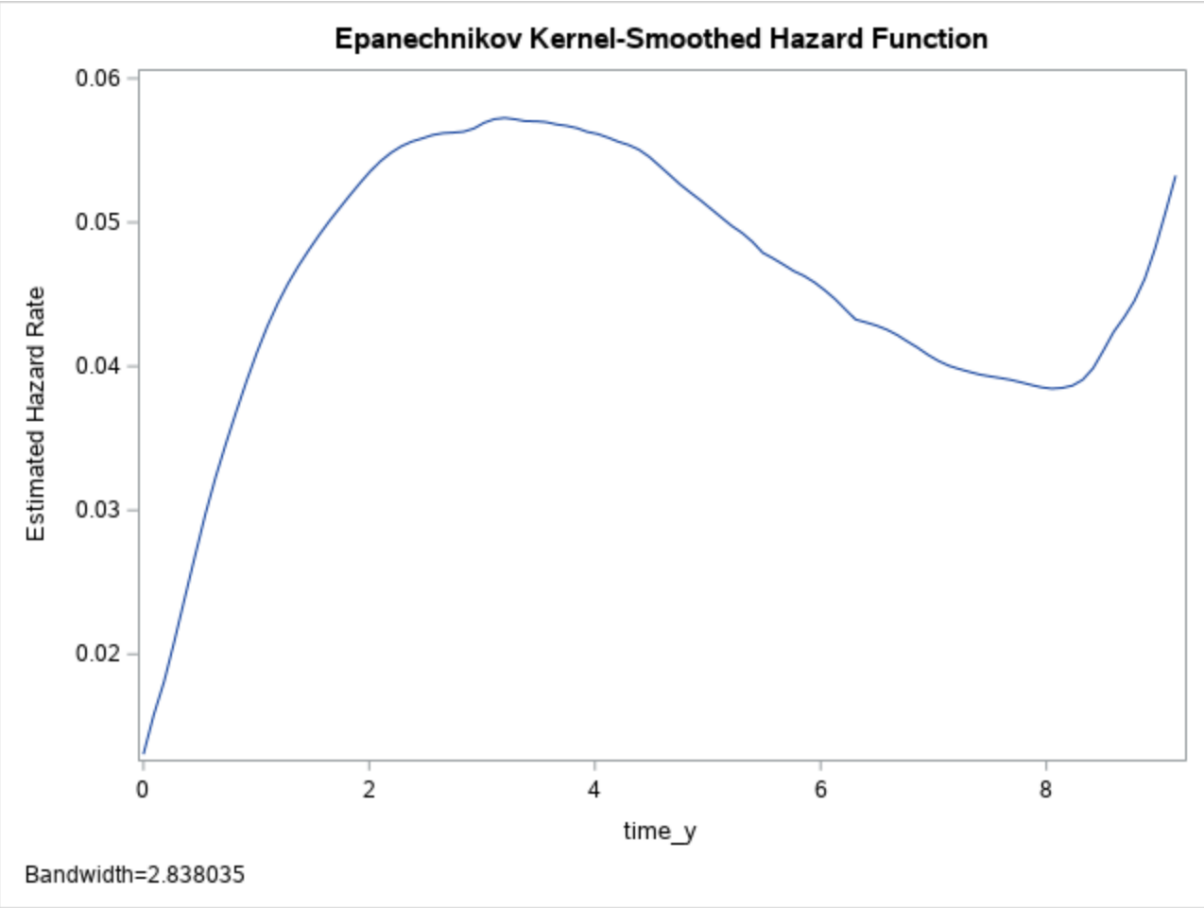

Figure C1

Kernel-smoothed estimate of hazard rate

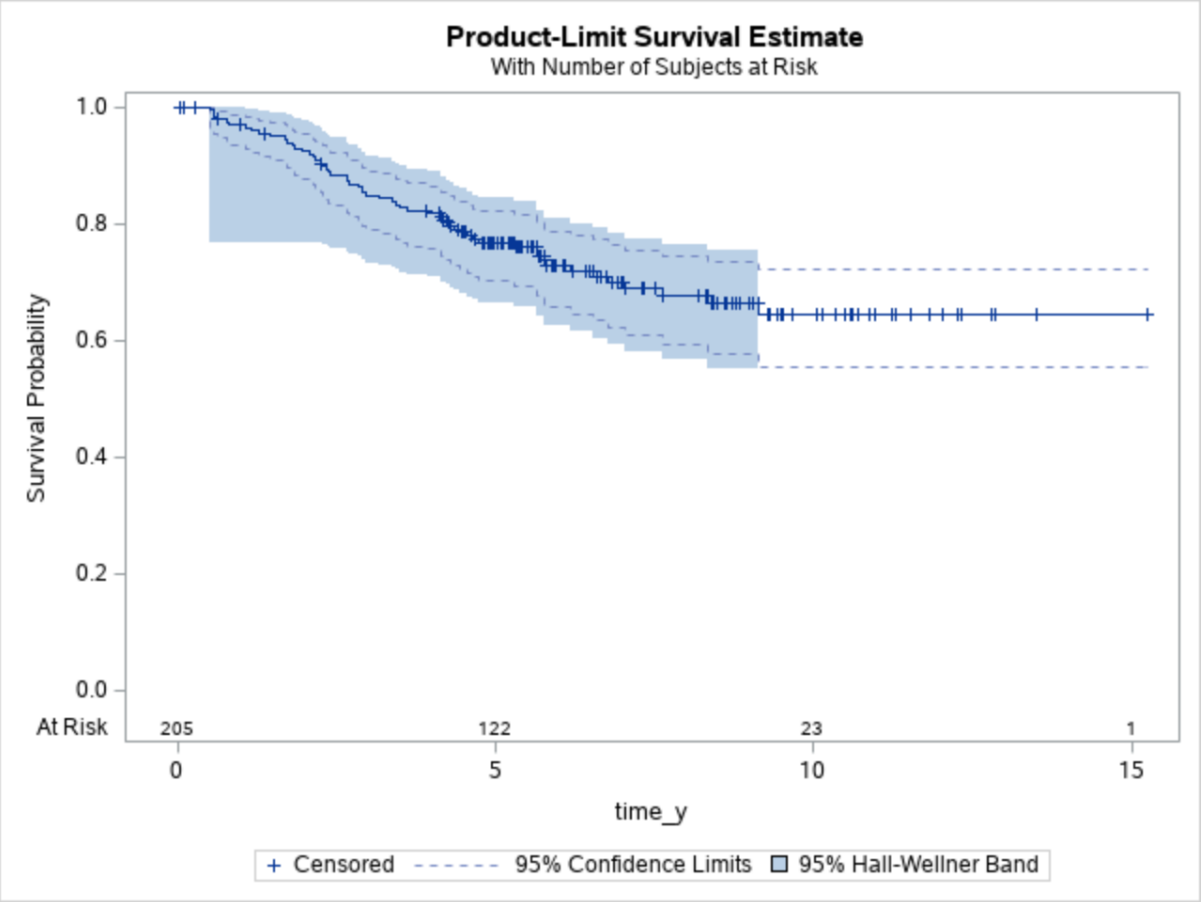

Figure C2

Overall survival Probability with 95% confidence band and pointwise confidence interval

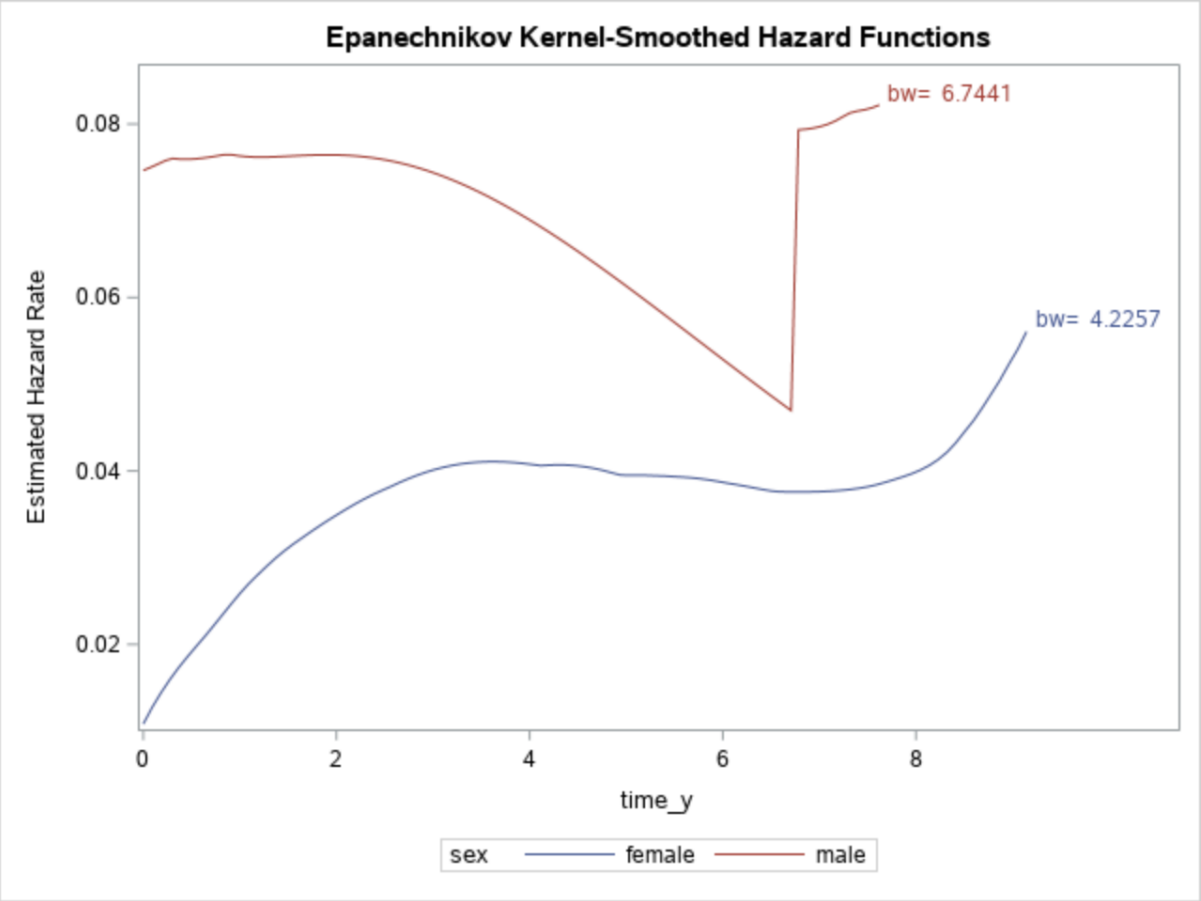

Figure C3

Kernel-smoothed estimate of hazard rate of two genders

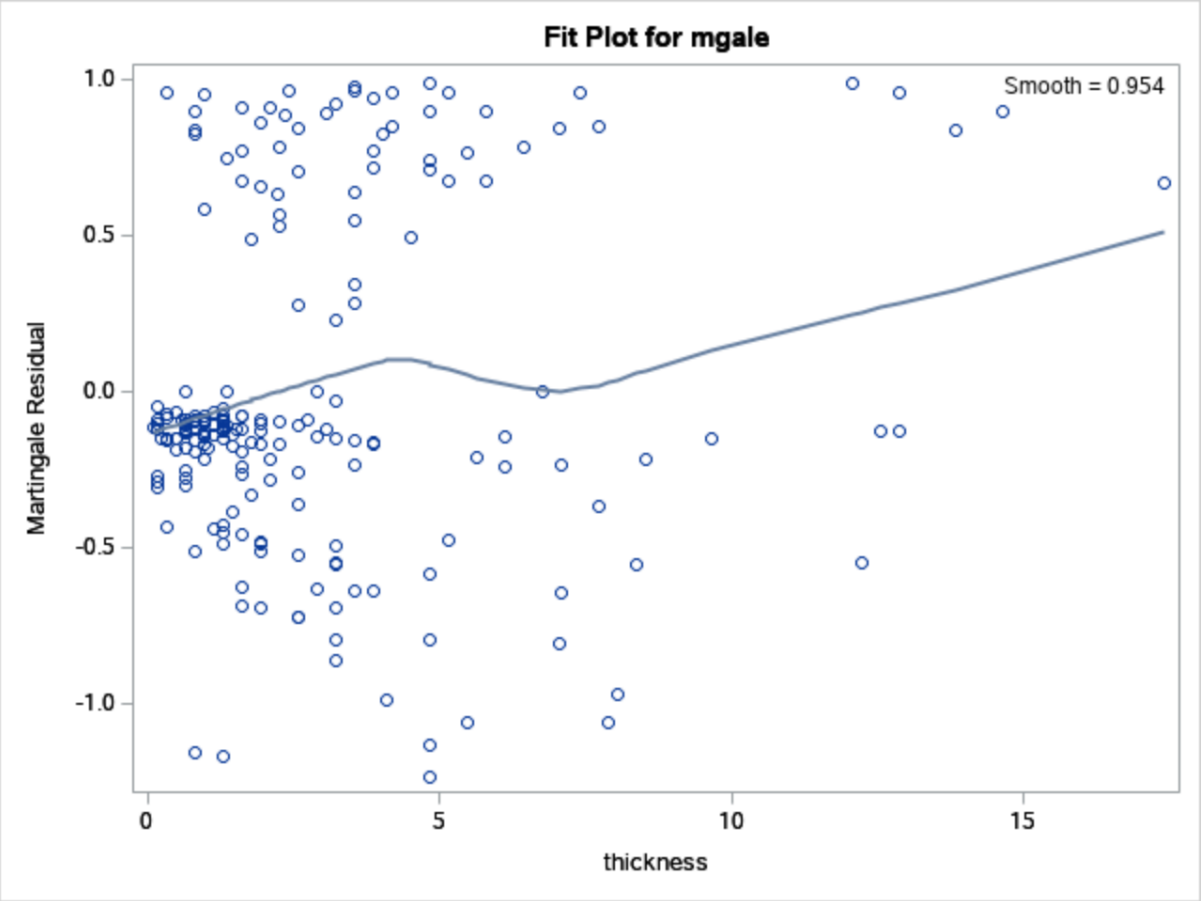

Figure C4

Martingale residual plot of thickness afte adjusted for age, sex, and ulcer

Figure C5

Martingale residual plot of age after adjusted for thickness, sex, and ulcer

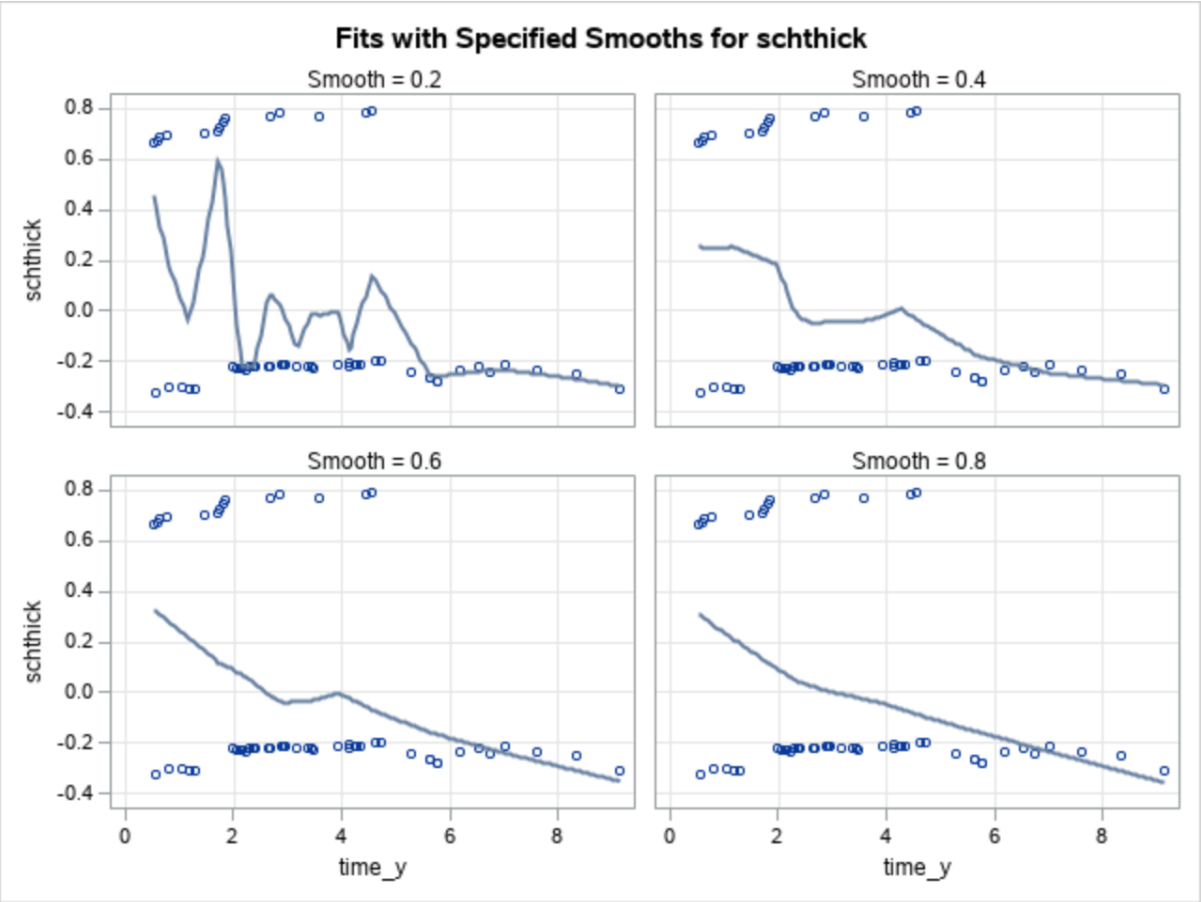

Figure C6

Schoenfeld residual plot for thickness of tumor

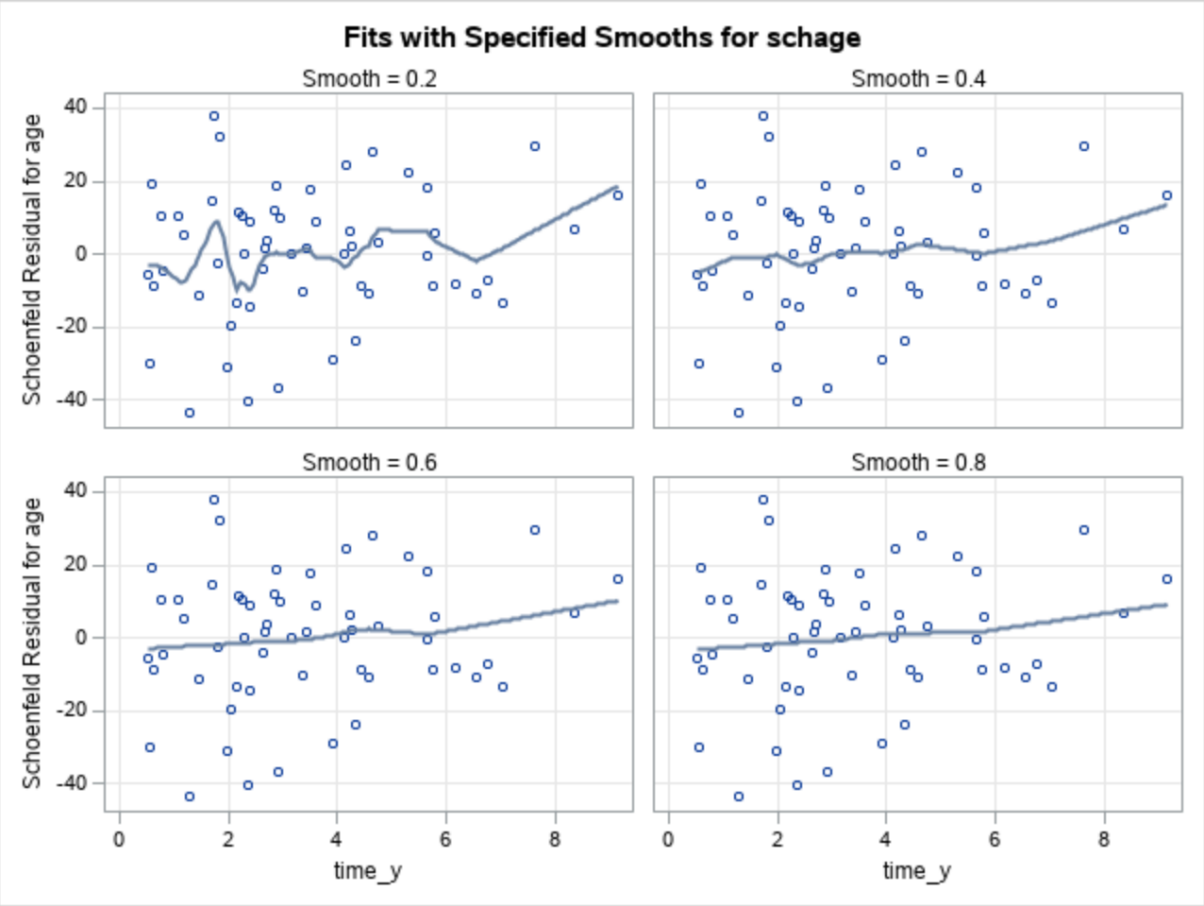

Figure C7

Schoenfeld residual plot for age

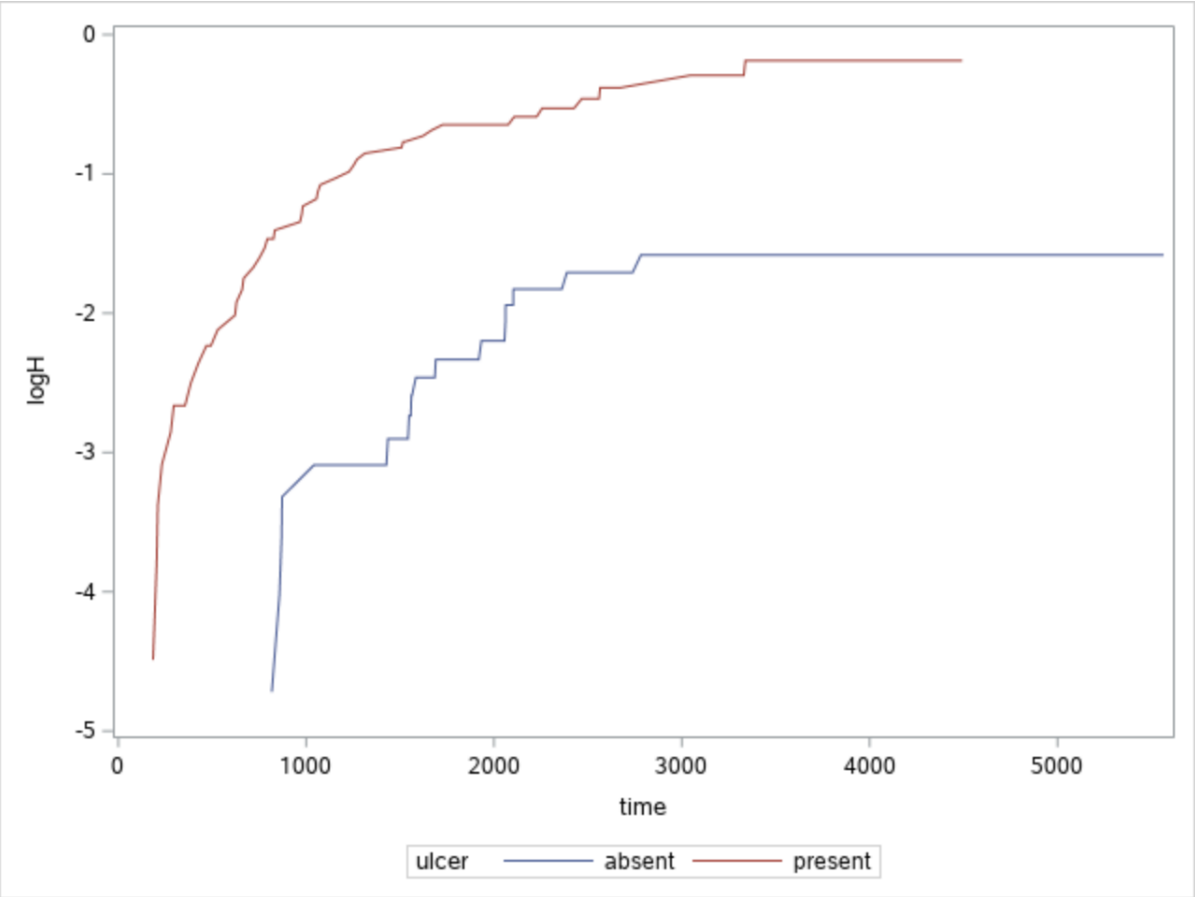

Figure C8

Log-log survival plot based on ulcer

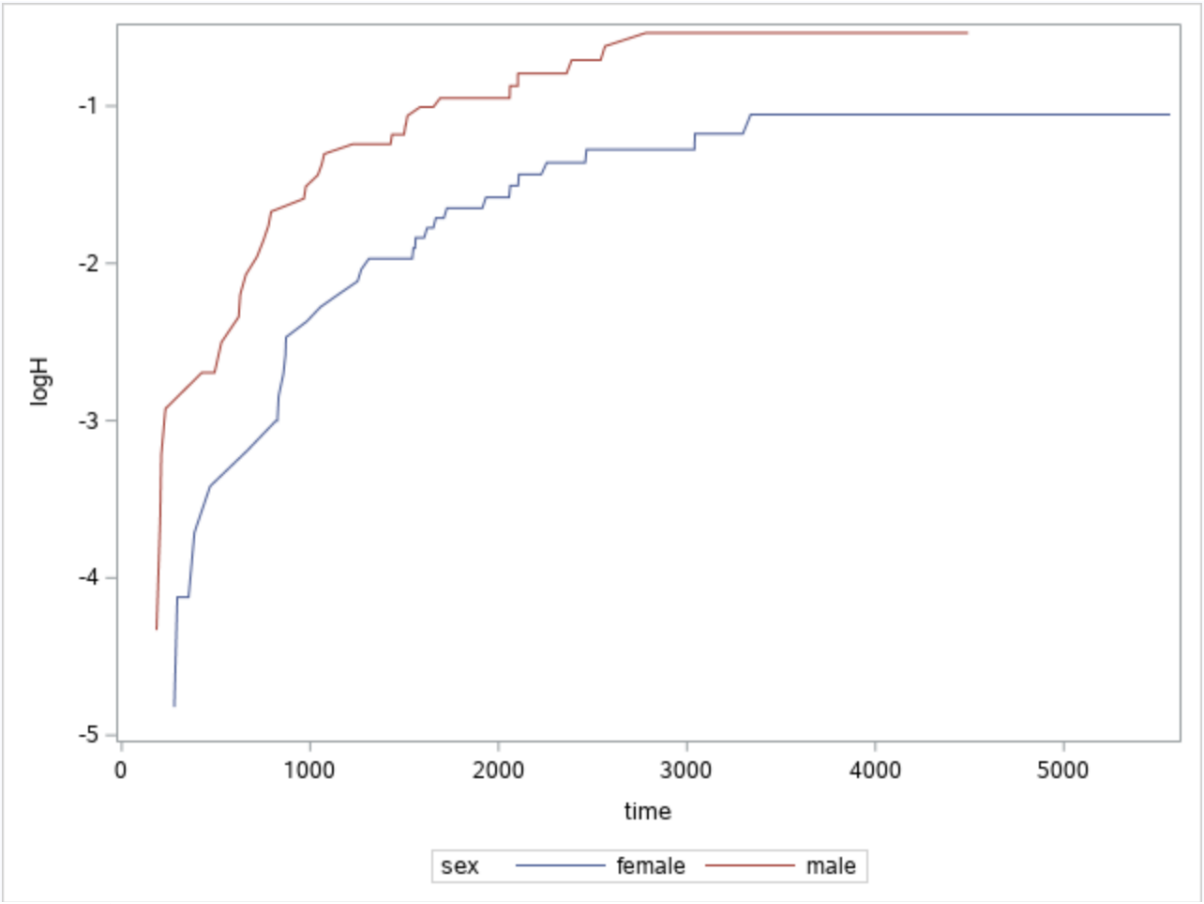

Figure C9

Log-log survival plot based on sex

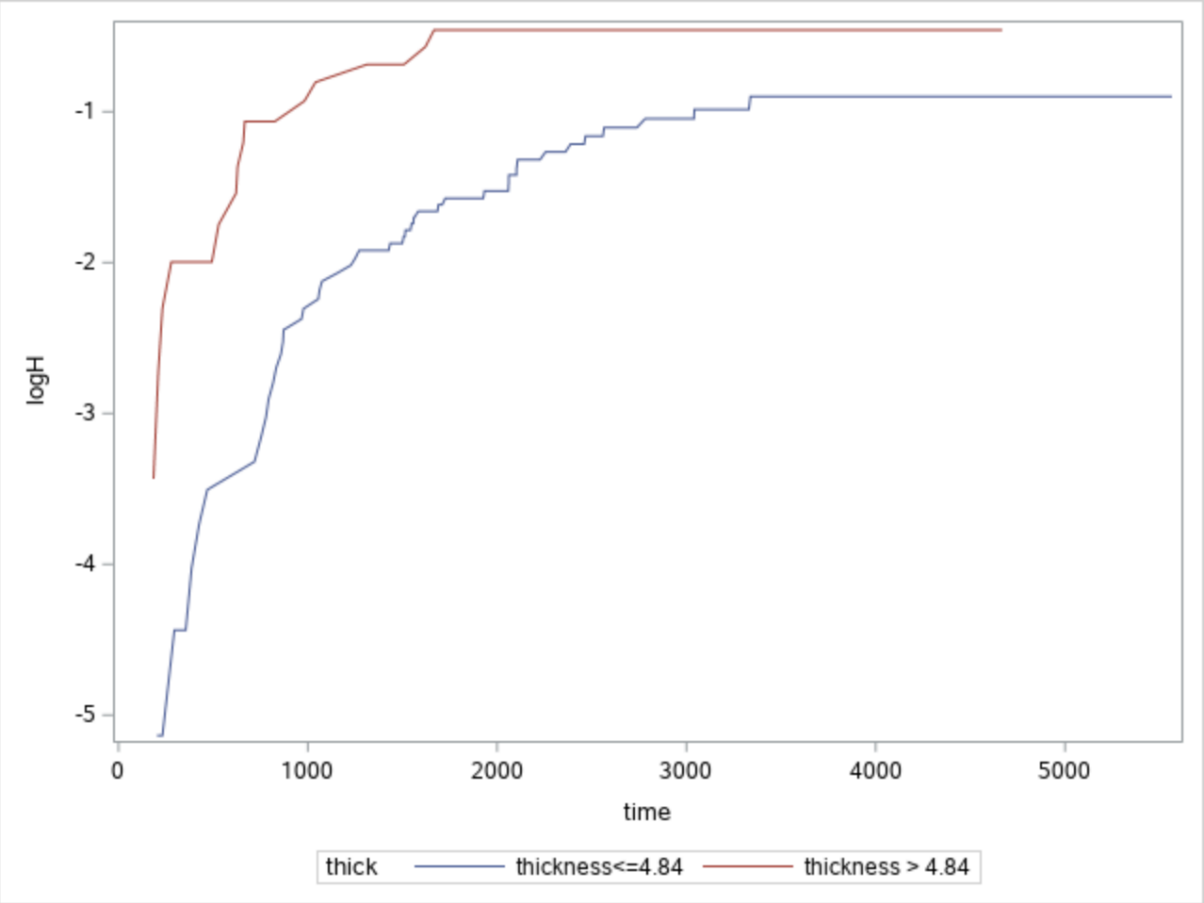

Figure C10

Log-log survival plot based on thickness level of tumor (two levels)

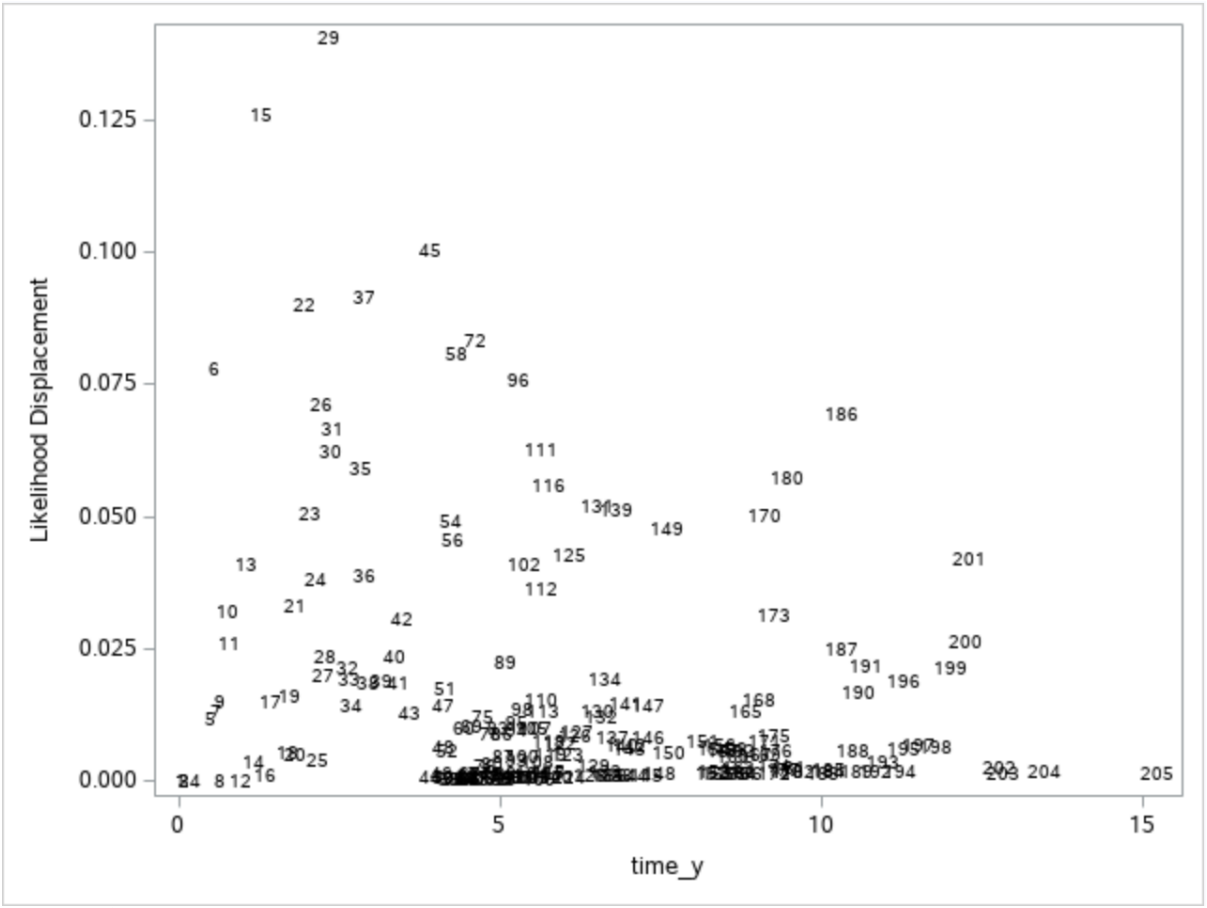

Figure C11

Plot of likelihood displacement scores to assess influence of the overall model

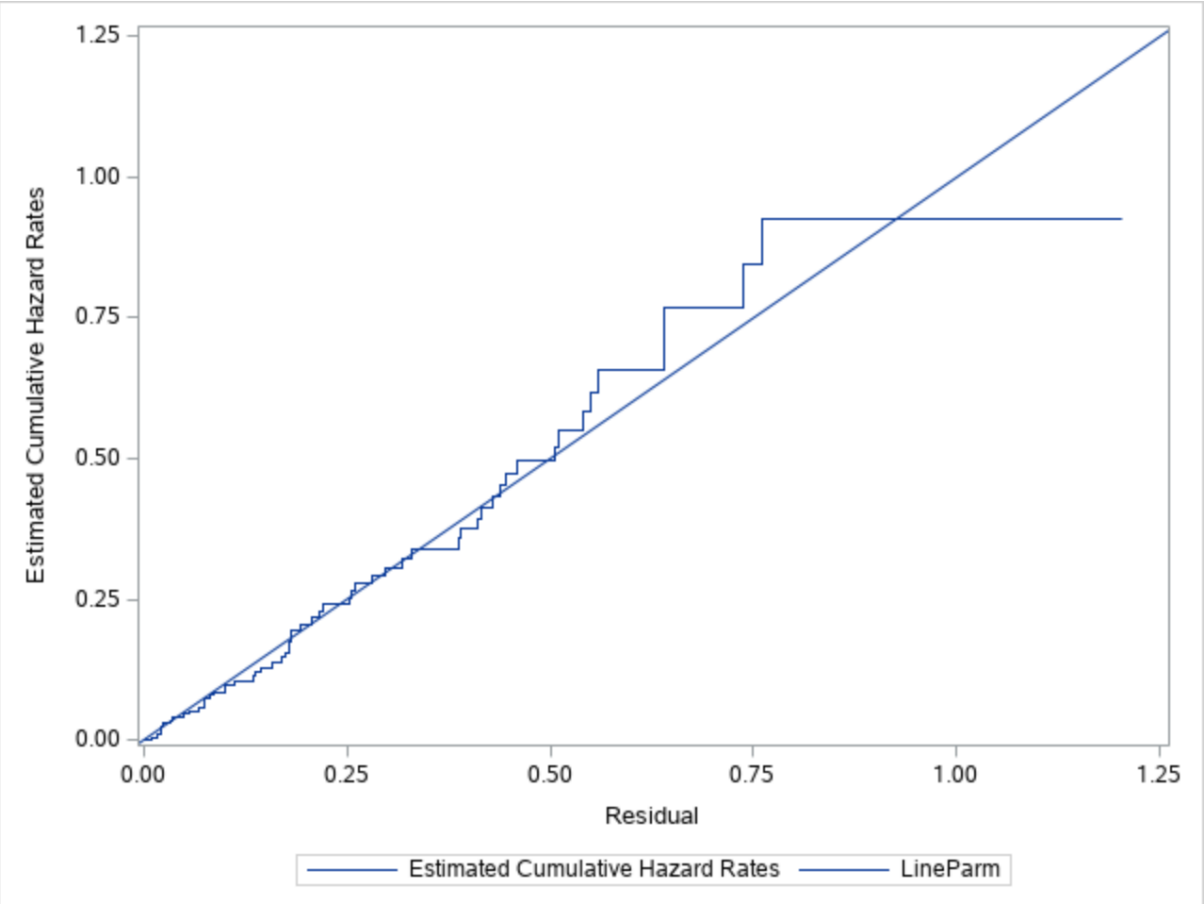

Figure C12

Cox-Snell Residuals of Final Model