Studying Alcoholism Relapse Time Using Survival Analysis

Abstract

The objective of this study is to identify what factors, or combination of factors, contributed to the relapse time of alcoholics. We are also interested in which of the two treatments (cognitive-behavioral or social) is superior in elongating relapse time compared to the other. We reviewed a total of 576 alcoholic patients and used a Cox PH model to analyze a range of genetic and environmental factors.

We concluded that family support, prior number of alcohol treatments, type of treatment, age, and the interaction between age and prior number of treatments were significant factors that influenced relapse time. Specifically, stronger family support, more prior treatments, younger ages were associated with a higher hazard of relapse to alcoholism. We also found the treatment emphasizing cognitive behavioral therapy improved relapse time more than social therapy.

Introduction

Alcoholism refers to excessive alcohol consumption. According to the National Institute of Alcohol Abuse and Alcoholism, this is an especially important issue in the United States where nearly 88,000 people die from alcohol-related causes annually (making alcohol the fourth leading preventable cause of death in the United States). Alcoholics:

- have a shorter life expectancy by about 10-12 years.
- play a major role in more than half of all automobile fatalities.
- make up a quarter of all people who commit suicide.

According to *Alcohol MD*, approximately 90% of alcoholics experience at least one relapse in the four years following treatments. As of January 2016, it was estimated that there are 60,000 Alcoholics Anonymous groups and 1.2 million members along with 1,450 correctional facilities in the US alone (According to the official Alcoholics Anonymous Website). Because alcoholism is such a dangerous issue that affects so many, it is critical for us to study the relapse time for alcoholism.

The relapse time from alcoholism can be affected by a variety of genetic and environmental factors. Some of these include depression, smoking, cognitive function, family support, number of prior alcoholic treatments, age, race, type of treatment and quality of clinic. In this study, we will be asking the following research question: **How do the previously mentioned factors, or**

combination of factors, affect relapse time? We will look especially at **how social support treatment and cognitive behavioral therapy differ in affecting a patient's relapse time.**

The rest of this report is organized in the following way:

I. **Description of Data & Subjects**: Investigation of data and subjects, making appropriate adjustments, and preparing it for our empirical analysis
II. **Univariate analysis**: Reviewing the characteristics of sample population by each of the variables, and understanding the singular effect of each predictor on the relapse time
III. **Fitting the Model**: Description of the model fitting process and the models obtained
IV. **Assessment of Model Adequacy**: Looking at the Cox PH assumption, Influential Points, and Goodness of Fit
V. **Results & Discussion**: Presentation of the result and final discussion on the topic

I. Description of Data & Subjects

The data in this study were collected from 576 patients of two different clinics, one in California and one in New York. We have access to the following data for each patient:

Variable Name	Description	Coding/Unit
relapse	Event indicator of relapse	1 = yes (subject relapsed to drinking) 0 = censored (did not relapse)
r_time	Time to drinking relapse	number of days
depression	Measure of depression ranging	from 0 to 63 (lower scores correspond to less depression)
cig_mari	Smoking/Marijuana use in 3 months prior to admission	1 = both cigarette smoking and marijuana use 2 = cigarette smoking only 3 = marijuana use only 4 = neither
cognit	Measure of cognitive functioning	Score (higher scores correspond to higher functioning)
family	Degree of family support	1 = minimal, 2 = moderate, 3 = strong
prior	Number of prior alcohol treatments	Number of treatments
age	Age at Enrollment	In years
race	Subject's race	0 = white, 1 = non-white
trt	Treatment group	0 = treatment emphasizing social support and talk therapy, 1 = treatment emphasizing cognitive-behavioral therapy

clinic	Clinic variable	0 = California clinic, 1 = New York clinic

To clean the data, we ran histograms and boxplots for each variable to visualize any outliers or unusual subjects. When the outliers were clearly a data entry error, we deleted the observation. An example of this was deleting any observations that had a negative age. When outliers were not necessarily a data entry error, we kept the observation. Variables that had data entry errors were: *cognit*, *family*, *race*, *age*, and *cig_mari.*

Overall, the average relapse time for a patient in our study was 234 days (with standard deviation 8.27). Most subjects were white, had an average age of 35.75 and had 4.5 prior alcohol treatments. The majority of our subjects were from the in California. For more information on the mean and standard deviation of the variables that are continuous; or the proportions for variables that are categorical, see the univariate analysis section.

In our analysis, we treated *depression*, *cognit*, *prior*, and *age* as continuous variables and we treated *cig_mari*, *family*, *race*, *trt*, and *clinic* as categorical variables. **In the *cig_mari* categorical variable, we used *cig_mari*=4 as the baseline for more logical interpretation and we also converted family into a binary variable** (which will be explained further in the later section).

II. Univariate Survival Analysis

Here we are looking at the characteristics of the sample population and the "singular" effect of each variable on survival outcome using the Cox PH model. If variables are significant to survival time at the alpha = 0.25 level, we will consider them for our initial multivariate model.

Variable	**Description**	**Coefficient Estimate (β)**	**P-value**	**Hazard Ratio**
depression	mean=16.7 sd=9.65	.004	0.377	1.004
cig_mari	**1=both cig & mari** 18% **2=cig only** 18% **3=mari only** 30% **4=none** 34% (with 4 as baseline)	β_1 = 0.2134514	0.117	1.237
		β_2 = 0.2483492	0.063	1.281
		β_3= -.0715066	0.554	.930
cognit	mean=100 sd=0.42	.008	0.074	1.008
family	**1=minimal** 38% **2=moderate** 18% **3=strong** 43%	β_2=.1576579	0.246	1.170

	(with 1 as baseline)	β_3=.3723849	0.000	1.451
prior	Mean = 4.549195 Sd = 5.44908	.031	0.000	1.031
age	Mean = 35.75 Sd = 7.066251	-.013	0.054	.987
race	**1= nonwhite** 25% **0= white** 75%	-.253	0.022	.776
trt	**0= social support treatment** 47% **1= cognitive behavioral treatment** 53%	-.251	0.007	.777
clinic	**0= California Clinic** 70.3% **1= New York Clinic** 29.7%	-.181	0.078	.834

The only variables with p-values greater than alpha = 0.25 are *depression* and *cig_mari* (1=both cig & mari, 2=cig only). We determine *depression* is insignificant after performing the Wald test and the LR test(See Appendix for output). For *cig_mari*, because not all levels of *cig_mari* are insignificant, we will leave it for now and discuss it more later in the model fitting section.

III. Fitting The Model (Multivariate Analysis)

After cleaning the data and looking at the effect of predictors individually, we will fit our multivariate model using the Cox Proportional Hazard Model. There are 5 main steps taken in order to obtain our Preliminary Final Model, which we will discuss in the remaining of this section.

<u>1. Fitting a multivariate model containing all predictors significant at 20% level.</u>

Our initial multivariable model includes 8 predictors (all predictors in the data except *depression*): *Cig_mari, cognit, family, prior, age, race, trt,* and *clinic.* Since *cig_mari* and *family* are categorical with several levels, we use the indicator function to represent the coefficients, as shown in the table.

No. of subjects =	538	Number of obs =	538
No. of failures =	435		
Time at risk =	128557		
		LR chi2(11) =	50.02
Log likelihood =	-2445.2452	Prob > chi2 =	0.0000

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cig_mari						
1	.9908265	.1642828	-0.06	0.956	.7159211	1.371292
2	1.001125	.1628296	0.01	0.994	.7278521	1.376998
3	.9158555	.1144149	-0.70	0.482	.7169504	1.169943
cognit	1.006752	.0047483	1.43	0.154	.9974883	1.016102
family						
2	1.248459	.1820958	1.52	0.128	.9380398	1.661604
3	1.395204	.2165626	2.15	0.032	1.029235	1.891301
prior	1.034191	.0090449	3.84	0.000	1.016614	1.052072
age	.9686558	.0074771	-4.13	0.000	.9541113	.9834219
race	.8590452	.1037189	-1.26	0.208	.6780225	1.088398
trt	.7774088	.0759966	-2.58	0.010	.6418583	.9415855
clinic	.8792266	.099063	-1.14	0.253	.7050102	1.096494

2. Use a Backward Stepwise Selection method to eliminate the insignificant predictors.

After designing this **Initial Multivariable Model**, we observed that the variables *cig_mari*, *cognit, family2, race*, and *clinic* were insignificant. We then performed backward stepwise selection to eliminate the insignificant predictors from the initial model one by one, and obtained a model with all predictors having p-values smaller than 0.05.

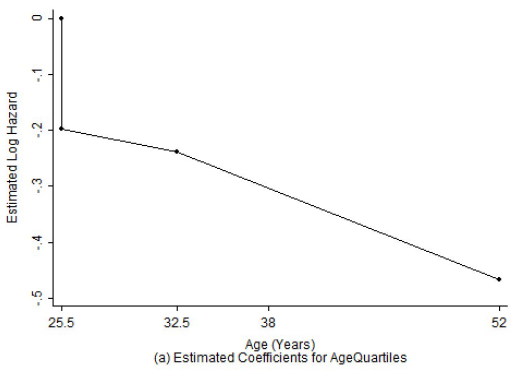
An important change to note is that when considering the predictor *family*, a categorical predictor with 3 levels, only *family3* was significant at the 5% level. In other words, only strong family support has a significant effect on relapse time whereas moderate family support (*family2*) was not significantly different from our baseline (*family1*, minimum support). Because of this, we decided to combine *family1* and *family2* into one group and make it our new baseline. We found that with this new baseline, strong family support has a significant effect on relapse time. We named this new binary variable *familystrong*.

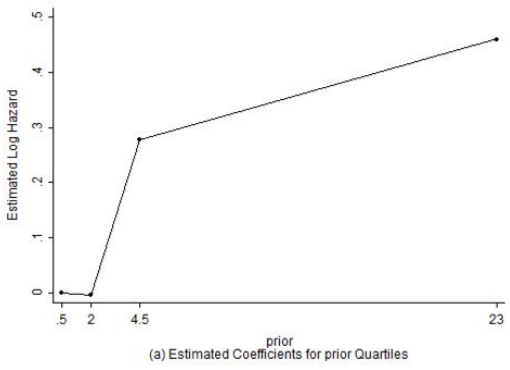
3. Checked if any of the predictors deleted in the previous step are confounders.

We inspected the predictors that we dropped to see if adding them back in would cause any large changes to the covariate estimates. We found that none of our dropped predictors were confounders. Therefore our model remained unchanged.

4. Look at the continuous predictors in our model (*prior* and *age*) to determine if any transformation was needed.

We plotted Log-Hazard against the quantile values of our continuous variables.





Based on the graph we weren't able to conclude whether or not the relationship was absolutely linear or not, so we performed a Frac Poly test. The Frac Poly test results informed us that no transformation was necessary (See Appendix for output).

5. Determine if there was any interaction between the significant predictors.

We tested all possible interactions in our model to see if they were significant and found one significant interaction between *prior* and *age*, which we named (*priorXage*). We included it in our **Preliminary final model**:

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
familystrong	.265544	.1017532	2.61	0.009	.0661115	.4649766
prior	.1431881	.0429941	3.33	0.001	.0589211	.2274551
age	-.0124558	.0092716	-1.34	0.179	-.0306278	.0057163
trt	-.2211978	.0951988	-2.32	0.020	-.4077841	-.0346116
priorXage	-.002846	.0011385	-2.50	0.012	-.0050775	-.0006145

IV. Assessment of Model Adequacy

After forming our Preliminary Final Model, we assessed its strength in the following 3 aspects:

1. Checked the Proportional Hazard Assumption

We used two methods to check the assumption. First we plotted the Schoenfeld residuals to see if the residuals are correlated with time. Since all of the plots showed no obvious pattern over time, the assumption is valid (output shown in Appendix). Secondly we used the *phtest*

function in stata to perform covariate specific test. Since none of the predictors had p-value smaller than 0.05, we confirmed that the PH assumption holds(output shown in Appendix).

2. Identified Influential Points and Poorly-fit subjects.

To identify the influential points, we checked the scaled score residuals. We looked at the likelihood displacement statistic versus the martingale residual plot provide us an overall summary of influential points (full output shown in Appendix). We were then able to identify these points in the plot. These influential points might potentially change our estimates, so we should further discuss them with the subject matter experts to refit the model appropriately.

3. Evaluated The Overall Goodness-of-fit of our Preliminary Final Model

We performed a goodness of fit test in Stata which compared observed and expected events for grouped quantiles of the data.

```
Score test                                    chi2(9)     =     7.730
                                              Prob > chi2 =     0.5615

Likelihood-ratio test                         LR chi2(9)  =     7.572
                                              Prob > chi2 =     0.5777

(Table  collapsed on quantiles of linear predictor)
```

Quantile of Risk	Observed	Expected	z	p-Norm	Observations
1	40	35.197	.81	.418	56
2	36	39.941	-.624	.533	56
3	45	36.81	1.35	.177	56
4	43	48.523	-.793	.428	57
5	43	46.621	-.53	.596	54
6	41	47.877	-.994	.32	56
7	51	47.174	.557	.577	57
8	48	53.147	-.706	.48	55
9	55	45.715	1.373	.17	59
10	50	50.995	-.139	.889	52
Total	452	452			558

Since the |Z| scores for each quantile is less than 1.96, the test tells us that the model has a good fit in all quantiles.

After assessing our model, we conclude that the **Preliminary Final Model** can sufficiently be our **Final Model.**

V. Results and Discussion

1. Empirical Results

For sake of clarity, we restate our **Final model**:

$$h(t,x) = h_0(t)\exp(\beta_f\, familystrong + \beta_p prior + \beta_a age + \beta_t trt + \beta_{pXa}\, prior * age)$$

where

$$familystrong = \begin{cases} 1 \text{ if strong family supprot} \\ 0 \text{ if minimum or moderate family support} \end{cases}$$

$prior$ = number of prior alcohol treatments

age = age at enrollment

$$trt = \begin{cases} 1 \text{ if cognitive\_behaviroal therapy} \\ 0 \text{ if social support and talk therpay} \end{cases}$$

$priorXage$ = interaction term between $prior$ and age

(where the coefficients can be found in the previous section).

The Hazard ratio, p-value, and Hazard Ratio Confidence Interval is shown below:

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
familystrong	1.30414	.1327004	2.61	0.009	1.068346	1.591977
prior	1.153947	.049613	3.33	0.001	1.060692	1.255401
age	.9876215	.0091569	-1.34	0.179	.9698364	1.005733
trt	.8015581	.0763074	-2.32	0.020	.6651225	.9659805
priorXage	.9971581	.0011353	-2.50	0.012	.9949354	.9993857

Based on our results:

- The hazard of relapse is 30% higher for patients with strong family support compared to those with moderate or low family support, holding other variables constant.
- The hazard of relapse is 19.8% lower for patients in cognitive behavioral treatment compared to those in social treatment, holding other variables constant.

Because of our *priorXage* interaction, the interpretation for the effects of *age* and *prior* treatment is a little bit different.

Specifically, the effect of one extra prior treatment on relapse time differs by age:
For example, holding other variables constant;

- For a patient at age 20, one additional prior treatment is associated with a 8.3% higher hazard.
- For a patient at age 30, one additional prior treatment is associated with a 5.1% higher hazard.
- For a patient at age 40, one additional prior treatment is associated with a 2.0% higher

hazard.

Similar, the effect of one extra year on relapse time differs by the number of prior treatment that the patient has had.
For example, holding other variables constant;

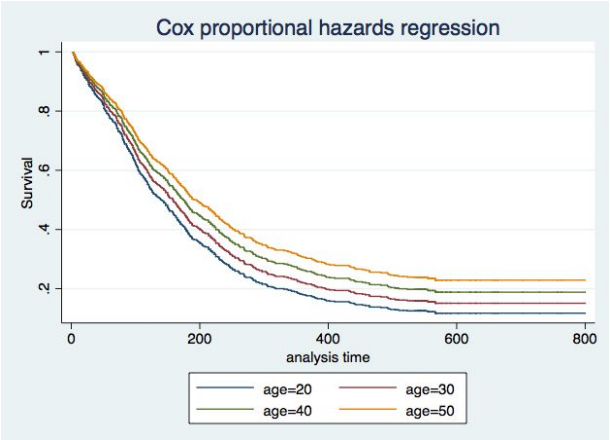
- For a patient that had 1 prior treatment, one additional year in age is associated with a 1.5% lower hazard.
- For a patient that had 5 prior treatments, one additional year in age is associated with a 2.5% lower hazard.
- For a patient that had 10 prior treatments, one additional year in age is associated with a 4.0% lower hazard.

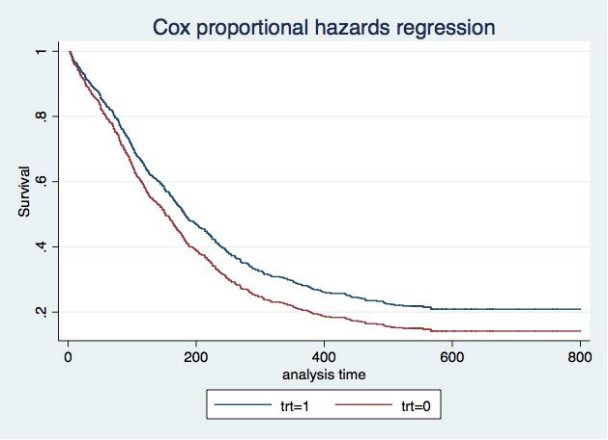
To better illustrate the different survival experiences for various sub-populations, we look at adjusted survival curves and the risk scores for patients at different percentiles.

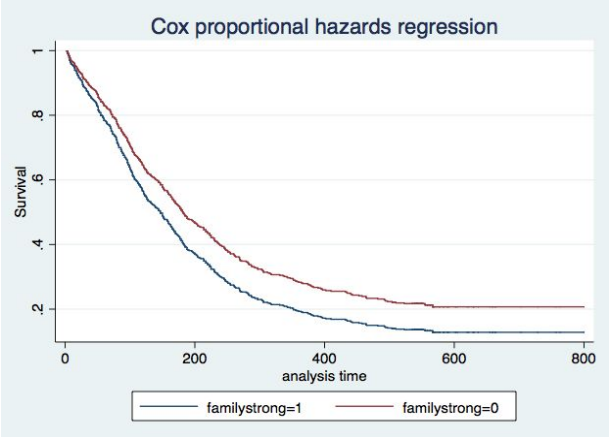
2. Adjusted Survival Curves

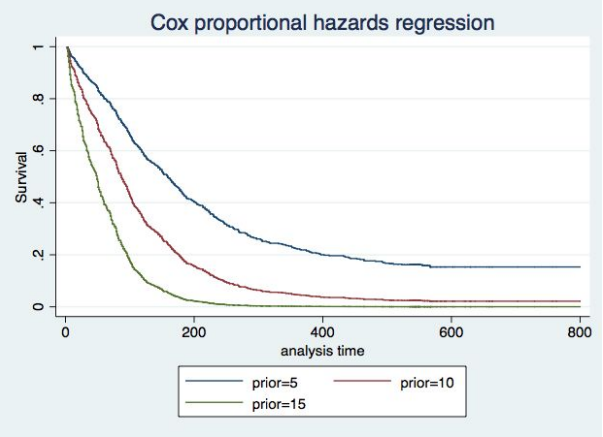
Here are two examples of how we compare the survival curves.

- Patient in the cognitive-behavioral treatment(*trt*=1) have a better survival experience (longer time to relapse) compared to those in the social treatment(*trt*=0).
- Older patients have longer relapse time compared to the younger patients (~50% of the patient at age 50 survived longer than 200 days versus only 35% of the patient at age 20 did).





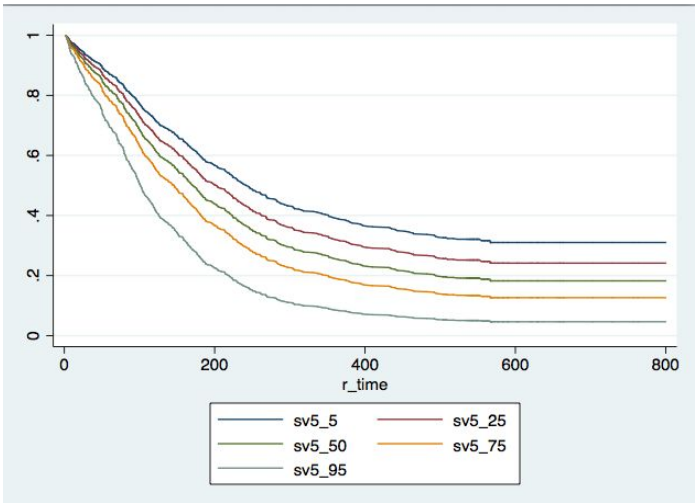




3. Percentile Risk Score Plot

The median survival time for patients with risk scores at the following percentiles are:

Risk Score Percentile	Survival Time (days)
5%	241
25%	203
50%	171
75%	147
95%	102



We see that it takes more than twice as long for a patient whose risk score is at 5 percentile to relapse to alcoholism compared to those with risk score at 95 percentile.

4.Conclusion

Overall, we conclude that the number of prior alcohol treatments, the age of the patient, the treatment type (cognitive-behavioral or social treatment), and the level of family support have a significant effect on a patient's relapse time for alcoholism. However, we did not find any strong evidence to show that depression, smoking history, cognitive score, race, and the patient's clinic had a significant effect. Also, we discovered that cognitive behavioral treatment was superior to social treatment in helping patient prevent relapse. The immediate next step would be to discuss these results with doctors and health professionals to ensure our findings make sense from a medical perspective.

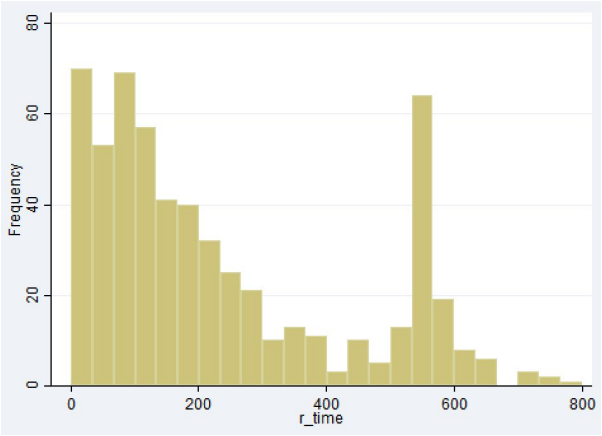
Possible further research could be performed to understand why more family support is associated with a higher hazard and why the cognitive behavioral treatment is better. Furthermore, we suggest that gender and income be considered as factors that could potentially affect relapse time to alcoholism.

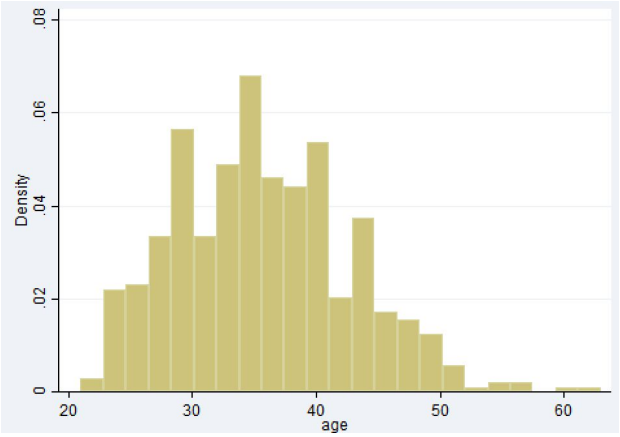
We hope the statistical findings from this study can further the our knowledge of alcoholism relapse time and improve the future care of alcoholism.

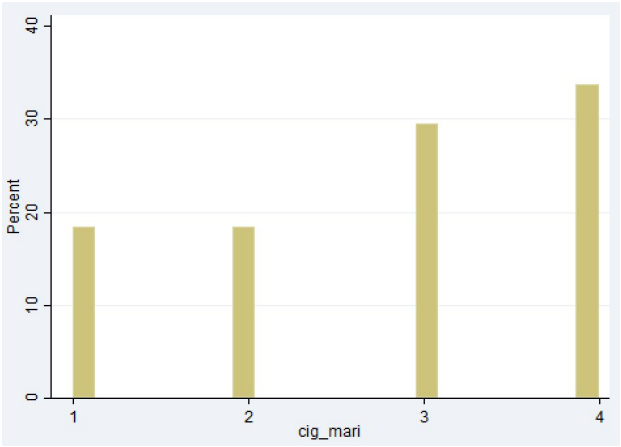
Appendix

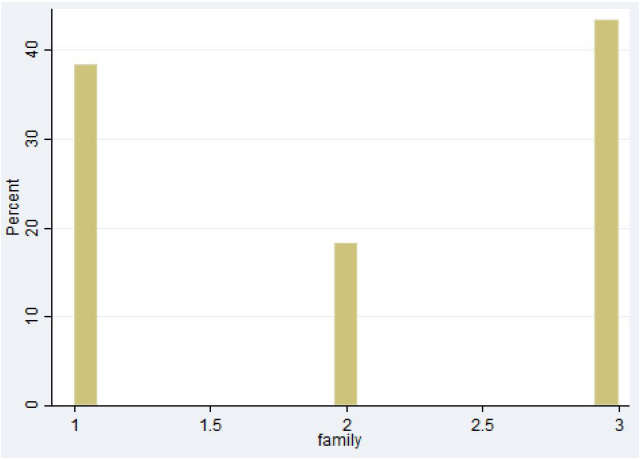
This appendix includes relevant plots and Stata output that we have mentioned throughout the paper. These outputs either directly support our findings in the paper, or facilitated our thinking.

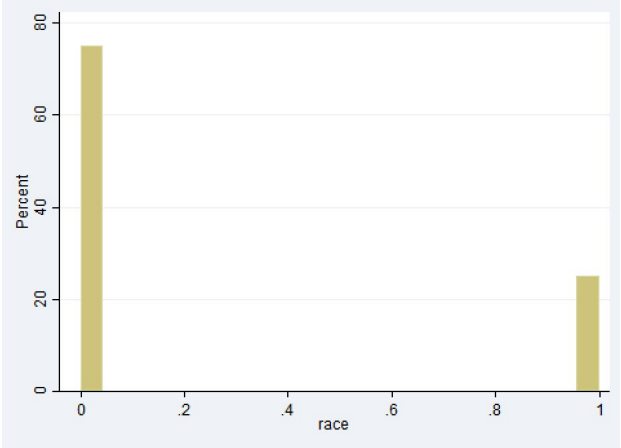
Histograms of Relapse Time and The Predictors

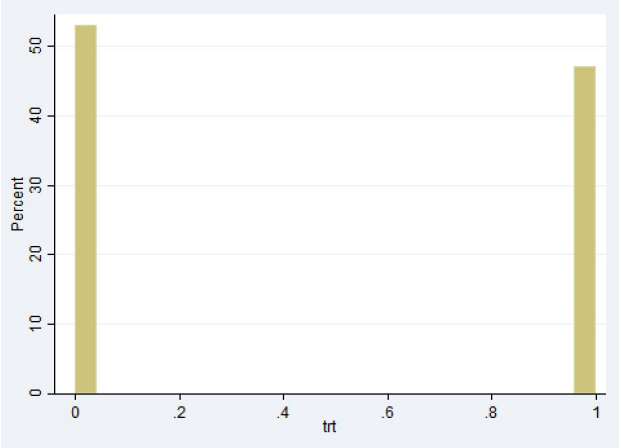


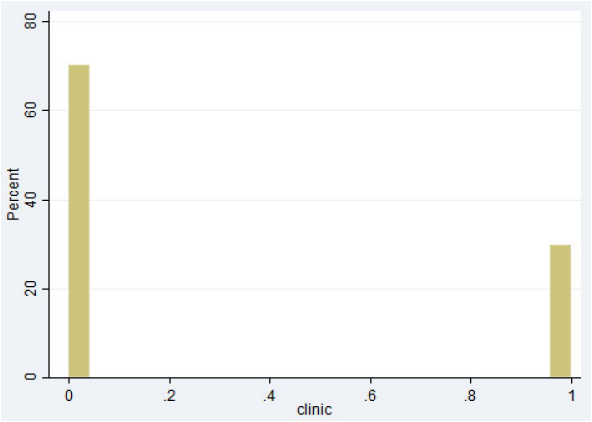


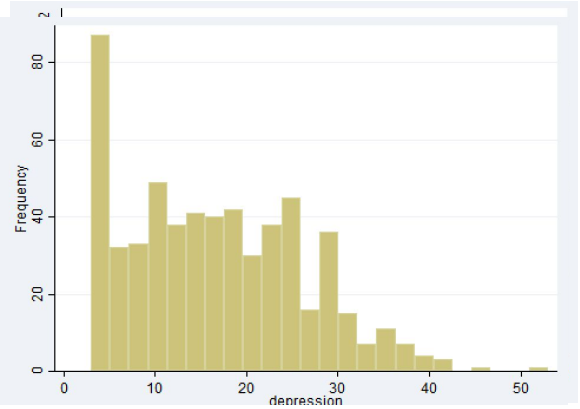












<u>To make sure that *depression* is not a significant predictor:</u>

Depression LR TEST:

```
No. of subjects =          547                 Number of obs    =         547
No. of failures =          442
Time at risk    =       131032
                                               LR chi2(11)      =       40.29
Log likelihood  =    -2497.512                 Prob > chi2      =      0.0000
```

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
depression	1.005601	.005135	1.09	0.274	.9955871	1.015716
cig_mari						
1	1.015085	.1660993	0.09	0.927	.7365794	1.398896
2	1.029264	.1656689	0.18	0.858	.7507898	1.411026
3	.9149115	.1139627	-0.71	0.475	.7167244	1.167901
cognit	1.008676	.0047527	1.83	0.067	.999404	1.018034
family						
2	1.283838	.185936	1.73	0.084	.9665671	1.70525
3	1.464758	.2214483	2.52	0.012	1.089125	1.969945
age	.9750328	.007255	-3.40	0.001	.9609165	.9893566
race	.8404169	.1001339	-1.46	0.145	.6653895	1.061485
trt	.7841094	.0760013	-2.51	0.012	.6484438	.9481586
clinic	.8725639	.0976524	-1.22	0.223	.7007057	1.086573

```
No. of subjects =          538                 Number of obs    =         538
No. of failures =          435
Time at risk    =       128557
                                               LR chi2(11)      =       50.02
Log likelihood  =    -2445.2452                Prob > chi2      =      0.0000
```

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
cig_mari						
1	.9908265	.1642828	-0.06	0.956	.7159211	1.371292
2	1.001125	.1628296	0.01	0.994	.7278521	1.376998
3	.9158555	.1144149	-0.70	0.482	.7169504	1.169943
cognit	1.006752	.0047483	1.43	0.154	.9974883	1.016102
family						
2	1.248459	.1820958	1.52	0.128	.9380398	1.661604
3	1.395204	.2165626	2.15	0.032	1.029235	1.891301
prior	1.034191	.0090449	3.84	0.000	1.016614	1.052072
age	.9686558	.0074771	-4.13	0.000	.9541113	.9834219
race	.8590452	.1037189	-1.26	0.208	.6780225	1.088398
trt	.7774088	.0759966	-2.58	0.010	.6418583	.9415855
clinic	.8792266	.099063	-1.14	0.253	.7050102	1.096494

Because chi squared value is = 3.48

51.43-50.02=1.41 so the model with depression is not significantly different.

Tests Performed To Check Transformations

Fractional polynomial model comparisons:

prior	df	Deviance	Dev. dif.	P (*)	Powers
Not in model	0	5142.130	19.898	0.003	
Linear	1	5127.288	5.055	0.409	1
m = 1	2	5127.180	4.948	0.293	.5
m = 2	4	5122.706	0.474	0.789	-1 -1
m = 3	6	5122.232	—	—	-1 -1 3

(*) P-value from deviance difference comparing reported model with m = 3 model

Fractional polynomial model comparisons:

age	df	Deviance	Dev. dif.	P (*)	Powers
Not in model	0	5141.636	15.410	0.017	
Linear	1	5127.288	1.062	0.957	1
m = 1	2	5126.590	0.365	0.985	3
m = 2	4	5126.516	0.291	0.865	-2 3
m = 3	6	5126.225	—	—	.5 .5 1

Checking the Proportional Assumption

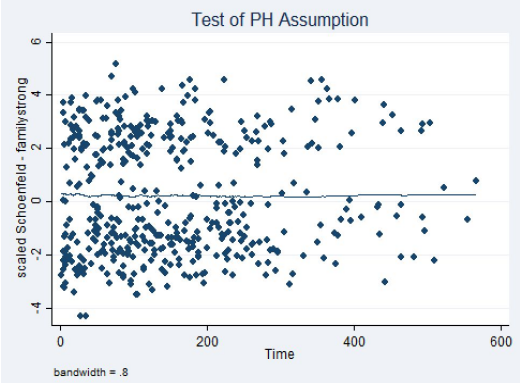
```
estat phtest, detail
```

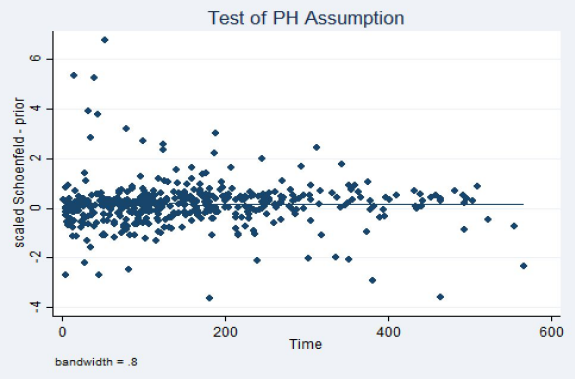
Test of proportional-hazards assumption

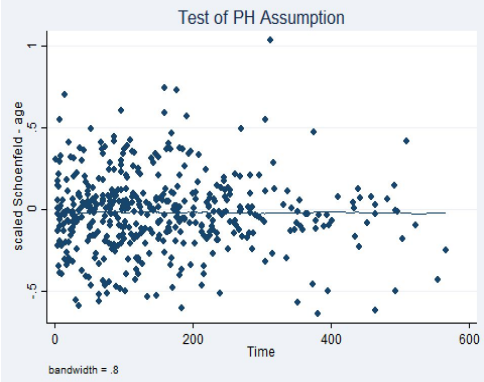
Time: Time

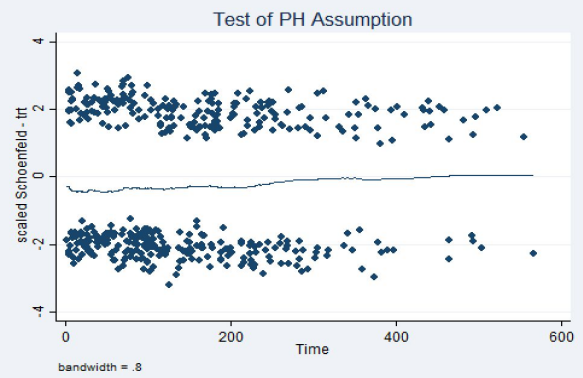
	rho	chi2	df	Prob>chi2
familystrong	0.01580	0.11	1	0.7375
prior	-0.01622	0.12	1	0.7312
age	-0.05570	1.42	1	0.2333
trt	0.08413	3.14	1	0.0763
priorXage	0.03751	0.63	1	0.4263
global test		9.98	5	0.0759

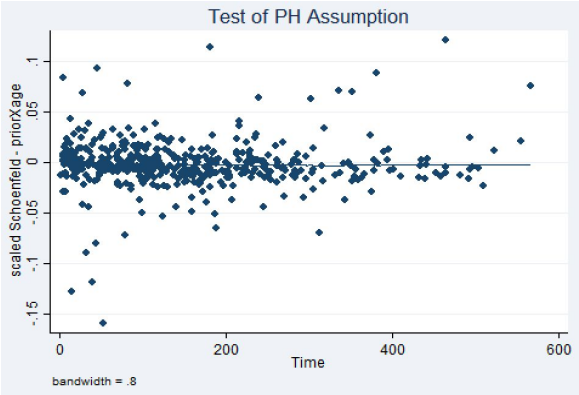
Since the p-values are all >0.05, the PH assumption is met.





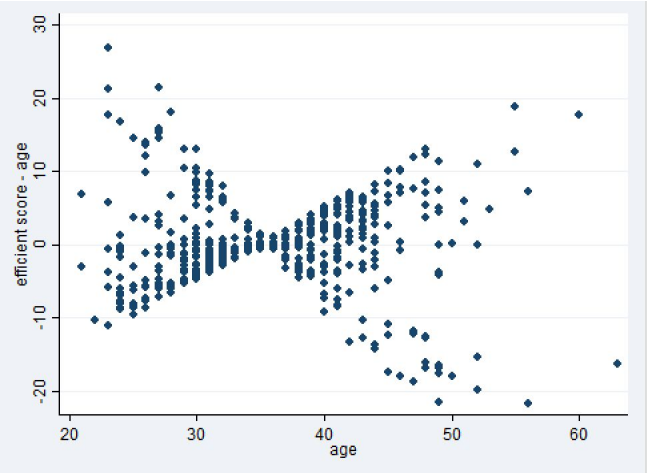


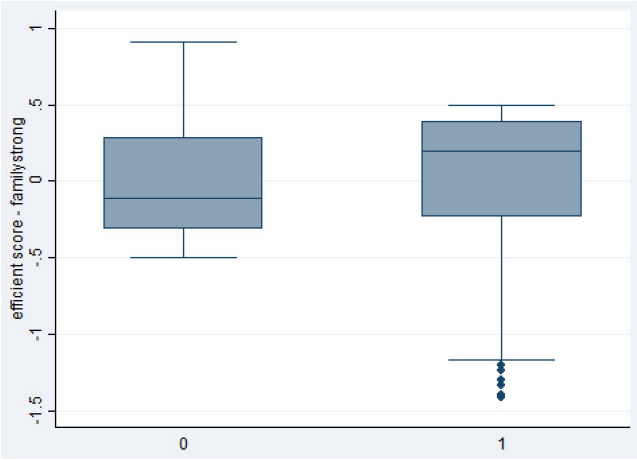


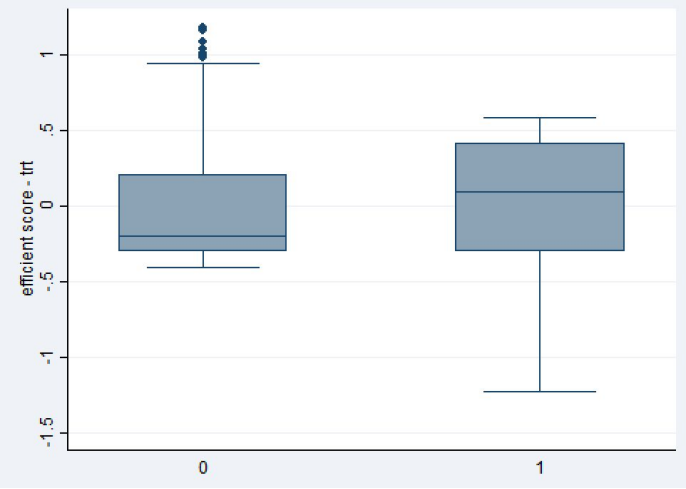


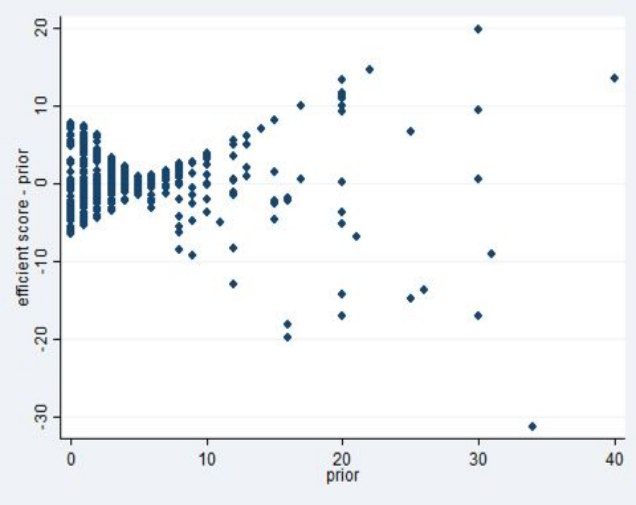
Identifying Influential Points

Score residual plots

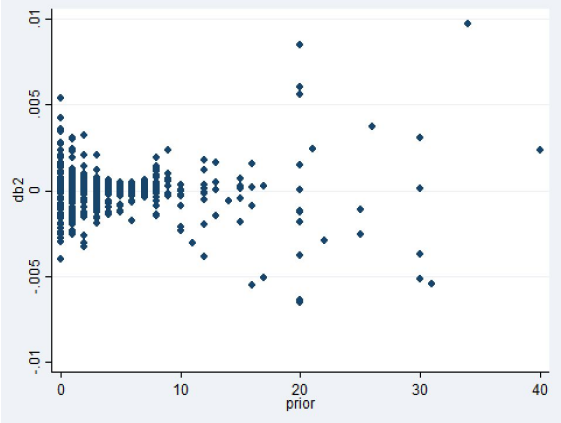


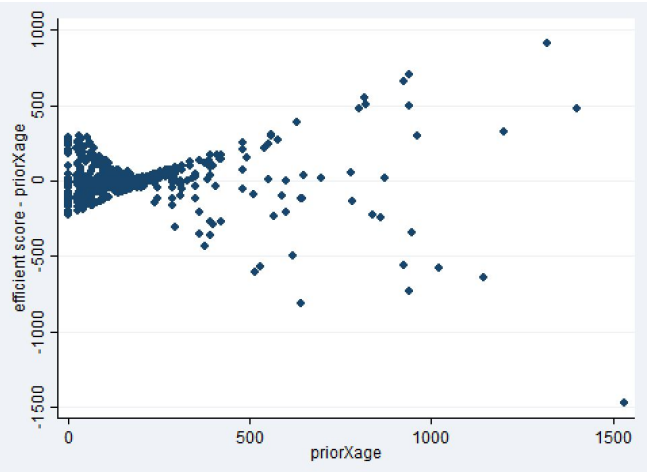


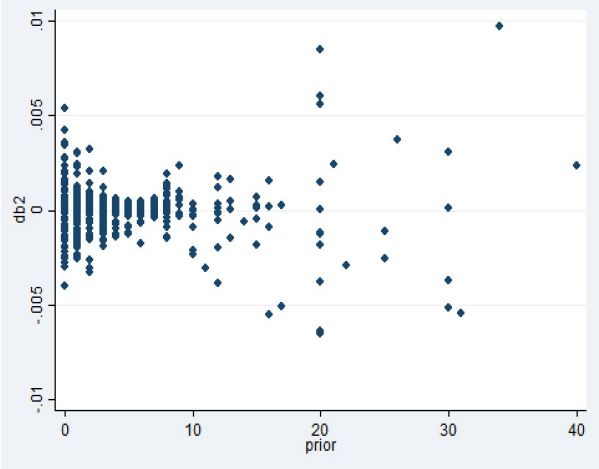


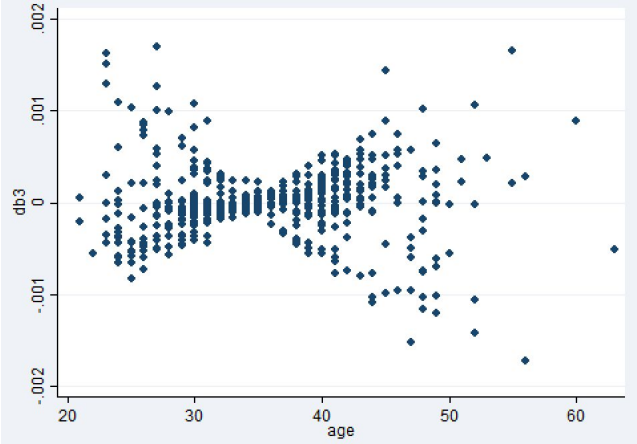


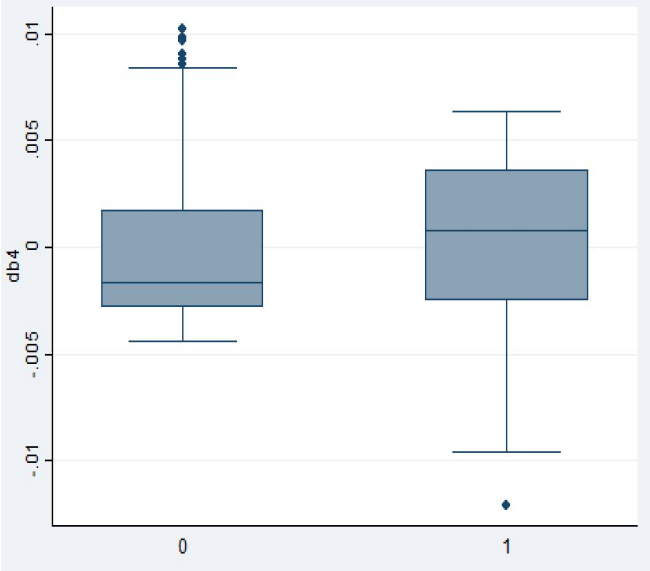
Scaled score residuals:

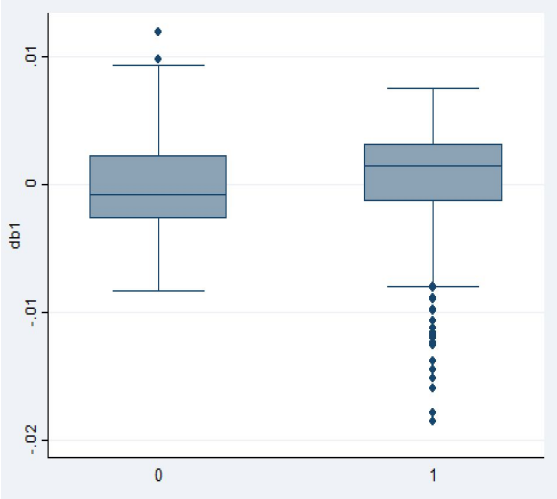




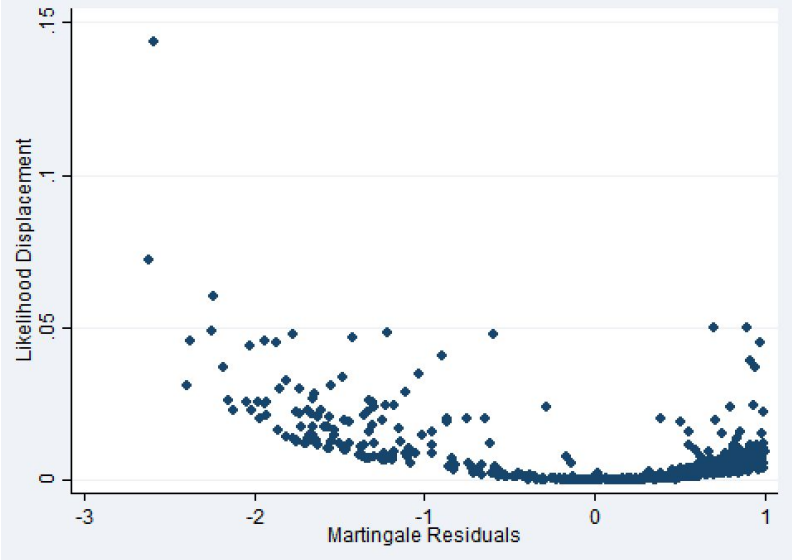

db2
.01
.005
0
-.005
-.01
0
10
20
30
40
prior


db3
.002
.001
0
-.001
-.002
20
30
40
50
60
age


db4
.01
.005
0
-.005
-.01
0
1


db1
.01
0
-.01
-.02
0
1

Likelihood displacement statistic versus the martingale residual plot provide us an overall summary of influential points. And we were able to clearly identify them in the plot.



Goodness of Fit Test

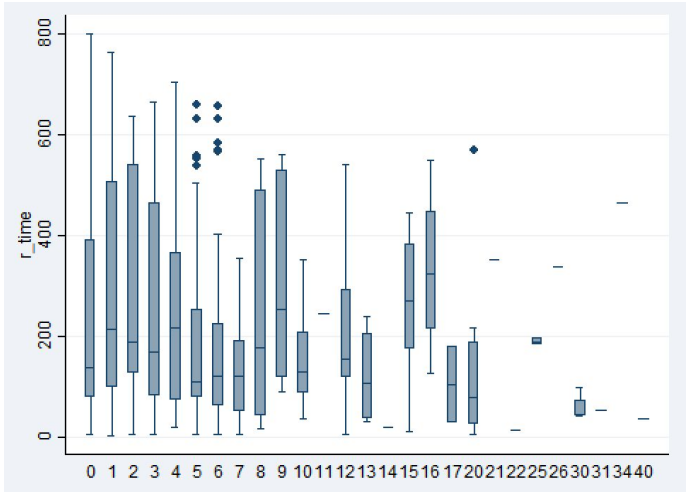
```
Score test                                  chi2(9)     =     7.730
                                            Prob > chi2 =     0.5615

Likelihood-ratio test                       LR chi2(9)  =     7.572
                                            Prob > chi2 =     0.5777
```

(Table collapsed on quantiles of linear predictor)

Quantile of Risk	Observed	Expected	z	p-Norm	Observations
1	40	35.197	.81	.418	56
2	36	39.941	-.624	.533	56
3	45	36.81	1.35	.177	56
4	43	48.523	-.793	.428	57
5	43	46.621	-.53	.596	54
6	41	47.877	-.994	.32	56
7	51	47.174	.557	.577	57
8	48	53.147	-.706	.48	55
9	55	45.715	1.373	.17	59
10	50	50.995	-.139	.889	52
Total	452	452			558

Since the |Z| score for all of the quantile is less than 1.96, it tells us that the model has a good fit in all the quantiles.

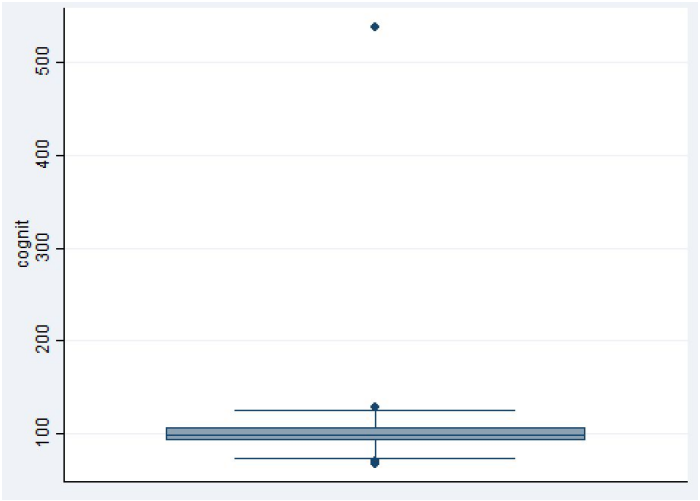
Prior



:

We considered splitting prior into two categorical variables, but after observing this graph, we can see that number of prior treatments does not have a strong correlation with r_time.

Cognit:



Because the outlier is so far from the other data, we assume this is an error and highly unlikely. Unless the data point represents Professor Smith's cognitive ability.

All possible interactions:

Interaction	P-value
familyStrong X prior	0.611
familyStrong X age	0.200
familyStrong X trt	0.378
prior X age	**0.013**
prior X trt	0.202
age X trt	0.260