

Survival Analysis Assignment – CHL5225H

Due date: October 24th, 2013

The data for this assignment is located at
www.andywillan.com/CHL5225H/data/brain.xls

The Excel spreadsheet “brain.xls” contains the data from a randomized trial of a new treatment vs standard of care for patients with malignant brain tumours. The primary outcome is survival time from the date of randomization.

The data set contains the following variables:

Treat: 1= new treatment 0 = standard of care

Resect75: 1 = >75% resection at reoperation, 0 = < 75% resection

Age: age in years at the start of treatment

Karn: 1 = Karnofsky performance score > 70, 0 = Karnofsky score < 70

Race: 1 = white, 0 = other races

Local: 1 = local radiation, 0 = whole brain radiation

Male: 1 = male, 0 = female

Nitro: 1 = previous exposure to nitrosoureas, 0 = no previous exposure

Weeks: time to death or end of follow-up

Event: 1 if death, 0 if alive at end of follow-up

Path: 1 = Glioblastoma, 2=Astrocytoma, 3=Other glial tumours, 4=Other

Grade: 1=active, 0 = inactive

NOTE:

- Do not include SAS code in your assignment.
- Do not exceed 20 pages (less is fine).

1. Plot Kaplan-Meier survival curves of time to death for the two treatment groups, and test for a treatment effect with the logrank test.
 - a. Estimate the median survival times for the two treatment groups.
 - b. Report the probability of death for each group at 26 weeks and 52 weeks of follow-up.
 - c. Test for a difference between the two groups with the logrank test.
 - d. Show the numbers of individuals at risk at quarterly intervals in the KM curve. Truncate the curve at 104 weeks.

2. Fit the semi-parametric Cox proportional hazards model of time to death, censoring individuals who were still alive at the last follow-up.
 - a. Using a graphical method, examine whether the proportional hazards assumption is appropriate for the treatment covariate.
 - b. Estimate the hazard ratio of death and 95% CI associated with treatment.
 - c. Identify other important predictors of time to death.
 - d. Fit a multivariable proportional hazards model and estimate the hazard ratio of death associated with treatment after adjusting for other covariates which are significantly associated with time to death.
 - e. Make a summary statement regarding the effectiveness of treatment on survival.

3. Investigate which parametric models would be appropriate to model this survival data. Discuss how you reached your conclusion regarding the best parametric model for this data. Using the most appropriate parametric model, estimate and interpret the effect of the treatment variable.