Assignment II ✎ (30%) DUE DATE: MAR 16, 2020

A retrospective cohort study of renal transplant was conducted. Databases were designed, which mainly consisted of two sections (i.e., enrollment and follow up). Enrollment contained baseline variables, of which some were fixed and some were varied over time. For instance, the fixed variables were operation date, birth date, sex, ABO group, HLA miss-matching, duration of dialysis before transplant. Time varying variables were follow up date, laboratory variables (e.g, serum creatinine, cholesterol, triglyceride, fasting blood sugar level, albumin, UA), medicine used and dosage (e.g., Cyclosporine dosage , Cardil), blood pressure and outcome.

The authors had treated infections (i.e., HBV, HCV, HIV) as fixed variables since baseline although they could be semi-time varying. For instance, infection could be negative test at baseline and later it was positive during follow up; and once positive (e.g. HBV antigen) should be positive for the rest of the study. Disease variables (i.e., diabetes, hypertension), were treated as semi-time varying regardless of fasting blood sugar or blood pressure. Two databases (i.e., enroll and follow databases) were side by side merged. Variables were labeled and described in detail in the databases.

1. Merge enrollment and follow up data:

    a) Enrollment data consist of fixed variables such as

Variable	Variable name
Date of operation	dateop
Age	age
Sex	sex
ABO group	bloodgr
Type of donor	typedonor

Type of HLA matching	hlamatch
HIV infection	hiv
HBS infection	hbs
HCV infection	antihcv
Mode of dialysis	mode

b) Categorized age as 3 groups based on distribution of data.

c) Follow up data consist of time varying variables such as

Variable	Variable name
No. Visit	visit
Date of visit	datevisit
DM diagnosis	diagdm
Hypertension	hyper
High cholesterol	highchol
Receiving immunosuppressive drug	
Cyclosporin	csa
Prednisolone	pred
MMF	mmf
Cardil	cardil
Graft failure	gf

2. Setting st data, in which the start and end dates were date of operation and date of visit, and failure was graft failure. After setting,

a) How many records were excluded?

b) What were the possible reasons of exclusion?

c) There were 43 records that showed “obs. begin on or after (first) failure”! Describe in detail how this happened and show the output.

d) Survival analysis

- Estimate the 2-5-years graft failure rates along with the 95% confidence interval
- What is overall graft failure rate and what does it mean?
- What is the median of graft failure and 95% CI (if applicable)
- Assess factors associated with graft failure step by step: Univariate and Multivariate analysis

Describe in detail how you have constructed the model, and what is the final model? Check assumptions for the final model, and interpret and report results.