Statistical Study on Mortality Among Prostate Cancer Patients

Sunil Kumar[1], Shivam Tiwari[2]

[1]Department of Statistics, R.B.S. College, Agra – 282002, India.

[2]Department of Statistics, University of Allahabad, Allahabad – 211002, India.

Email: drsunilpathak.rbs@gmail.com Mobile:9411002493

Abstract:

This project deals with the survival analysis of prostate cancer patients under different factors responsible for death. Prostate cancer is te second most common cause of male cancer deaths in western countries. The data set for this project has been taken from website of biostatistics. Survival analysis is one of the statistical methods for analyzing the data on time to event such as death, heart attack, recovery from surgery etc. In survival analysis, Kaplan - Meier method is a non-parametric method for obtaining survival estimates of respondent and it is specially used for censored observations for obtaining the survival estimate without treating them as missing observations. 'Competing risk analysis' refers to a special type of survival analysis that estimates the marginal probability of an event in presence of competing events. Cumulative Incidence Function is a method that provides more interpretable estimate for survival experience of multiple competing events for a group of subjects.

Keywords: Prostate cancer; Survival analysis; Competing risk analysis; Cumulative Incidence Function

1. Introduction:

Survival analysis is the one of primary statistical methods for analyzing data on time to event such as death, device failure, recovery, disease incidence etc. The basic data structure of survival analysis is tuple i.e. (start, stop, event), where the first two elements signifies the starting and ending time of study , while the third elements signifies whether the event has occurred or not .

Survival methods are only techniques capable of handling censored observations without treating them as missing data. Kaplan-Meier method is a non- parametric method for obtaining survival estimates of respondent and it is specially used for censored observations for obtaining the survival estimates without treating them as missing observations. In statistics, medical research, censoring is a condition in which the value of observation is partially known or we can say that the partial information about the observations are known as 'censoring'.

'Competing risk' analysis refers to a special type of survival analysis that estimate the marginal probability of an event in presence of 'competing events'. In survival data in

which each subject can experience only one several different types of events follows. In that case when only one of several different types of events can occur, we refer to probabilities of these events as 'competing risk'. A logical objective for 'competing risks' data is to assess the relationship of relevant variables to failure rate of any one of possible events allowing for competing risks of other ways to failure. We might also want to compare the hazard rate of one event with hazard rate of another event.

2. Literature Review:

Green and Byar (1980, Bulletin Cancer, Paris, 67, 477-488) have analyzed some data from a randomized clinical trial comparing treatment for patients with prostate cancer in stages 3 and 4. These authors assessed the effects of treatment on survival using an exponential regression model with treatment by covariate interactions. Treatment consisted of different doses of diethylstilbestrol (DES). Clinical considerations led to conclusions being drawn about the relationship between treatment and different causes of death. Green and Byar concluded that older patients and those with a history of cardiovascular disease were at higher risk of cardiovascular death and should therefore be given low-dose DES. Correspondingly, patients with more advanced disease, as measured by stage/grade category and tumor size, were at higher risk of cancer death and would therefore do better under high-dose DES.

3. Data source and Descriptions

We have used the dataset which has been collected by Green and Byar (1980, Bulletin, Paris, 67, 477-488) in a randomized clinical trail comparing treatments for patients with prostate cancer patients in 3 and 4 stage.The data set consists of 506 patients consist of 18 variables.

Description of variables: -

1. Panto -Number of patients
2. stage - Patient's stage of prostate cancer
3. rx - Different doses of diethylstilbestrol
4. Dtime -Months of follow-up
5. Status -Alive or different causes of death
6. Age -age of patient in years
7. wt - weight index=wt(kg)-ht(cm)+20
8. pf - patient's performance rating
9. hx - history of cardiovascular disease
10. sbp - systolic blood pressure/10dbp
11. dbp - Diastolic blood pressure/10
12. ekg - previous disease of patient
13. hg - serum hemoglobin (g/100 ml)
14. sg -combined index of stage and hist grade
15. ap -serum pro-static acid phosphates.
16. bm - bone metastases
17. sdate -date of study
18. sz - size of primary tumor
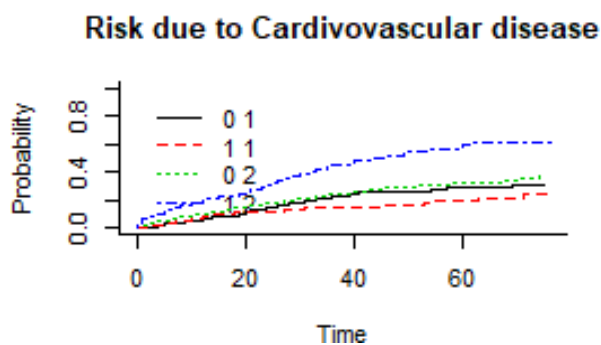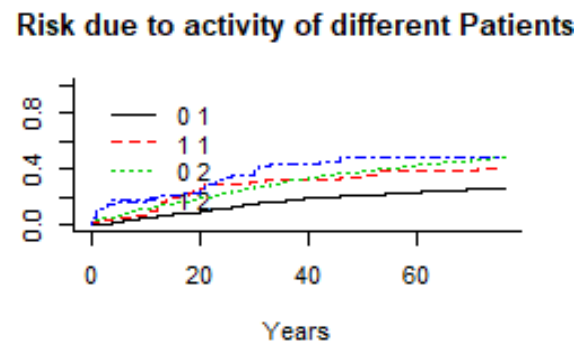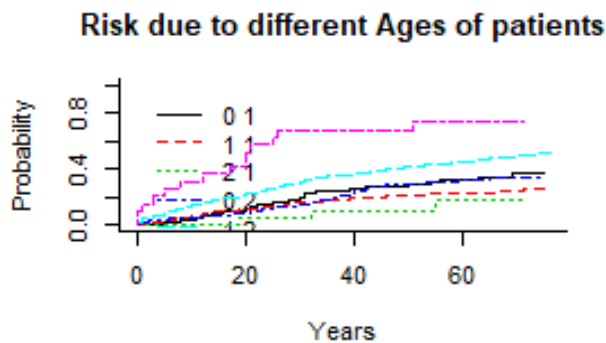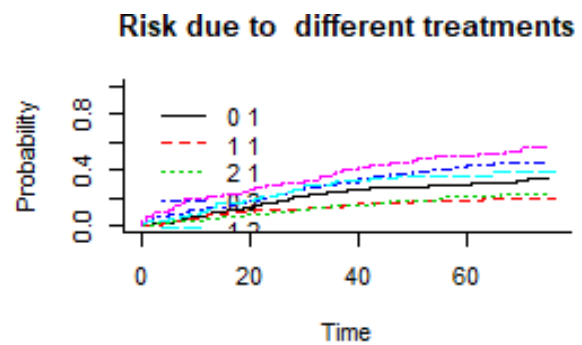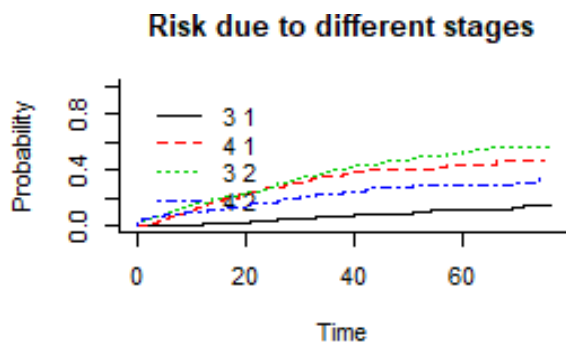
4. Methodology:

In survival analysis, Kaplan-Meier method is specially used to estimate the survival estimates in case of censored data set. It can be calculated by product of series of interval specific conditional probabilities. It can also be obtained from number of survivor at each observed survival time over total number of survivors at beginning of observations. Then we estimate the hazard rate at ordered time point where the event of interest occurs. The cumulative incidence for event at time is the cumulative sum up to time of these incidence values over all events failure times. The cause-specific approach of finding competing risks is to prefer a survival analysis for each event type separately where the 'competing event' are treated as censored categories. Cumulative Incidence Curve is derived from a cause-specific hazard function, provides estimates of marginal probability of an event in present of 'competing events'.

5. Results:

From the graph it can said that in prostate cancer any patient of stage 3 has less risk of mortality due to prostate cancer in comparison to risk of mortality due to other causes with the passage of time and risk of mortality of prostate cancer stage 4 patient is high in comparison to other stage and factor responsible for death. From the graph it can said that any patient of prostate cancer with passage of time has better probability of surviving, if he/she take 1 mg estrogen dose in comparison to 5 mg estrogen dose. Also, patients of placebo treatment have minimum probability of surviving and risk of mortality is high if patient take 5 mg estrogen dose and he/she is not suffering from prostate cancer. From the graph it can said that in prostate cancer the risk of mortality increases with the increase in age of patients. From the graph it can said that in prostate cancer, patient whose activity is normal has less risk of mortality in comparison to patient who take 50 "percent" bed rest during medication. From the graph it can said that in prostate cancer , patient who is free from CVD has less risk of mortality in comparison to patient who is suffering from CVD.

6. Interpretation:

From the above analysis it can be concluded that for prostate cancer patients stage 4 is more dangerous than stage 3, 1 mg estrogen dose is appropriate dose for patients , after a certain age the risk of mortality increases with increase in patient's age , risk of mortality is less when activity of patient is normal when it is more dangerous for patients having 50 per bed rest and also suffering from CVD.

Risk due to different stages

Risk due to different treatments

Risk due to different Ages of patients

Risk due to activity of different Patients

Risk due to Cardivovascular disease

References:

1. Kay, R. (1986). Treatment effects in competing-risks analysis of prostate cancer data. Biometrics, 203-211.
2. Kleinbaum, D. G. (1998). Survival analysis, a self-learning text. Biometri- cal Journal: Journal of Mathematical Methods in Biosciences, 40(1), 107- 108.
3. Klein, J. P., Van Houwelingen, H. C., Ibrahim, J. G.,&Scheike, T. H. (Eds.). (2016). Handbook of survival analysis. CRC Press.
4. Klein, J. P.,&Moeschberger, M. L. (2006). Survival analysis: techniques for censored and truncated data. Springer Science Business Media.
5. Liu, X. (2012). Survival analysis: models and applications. John Wiley& Sons.
6. Moore,D. F.(2016).Applied survival analysis using R. Switzerland: Springer.