ASSIGNMENT # 2

Submitted by

Julhas Sujan

Student ID: 2025166681

Course Name: Biostatistics-II

Submitted to

Dr. Ahmed Hossain

Professor, Department of Public Health

North South University

Email: ahmed.hossain@utoronto.ca

ahmed.hossain@northsouth.edu

Date of submission: 16 May 2021

Q1 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.

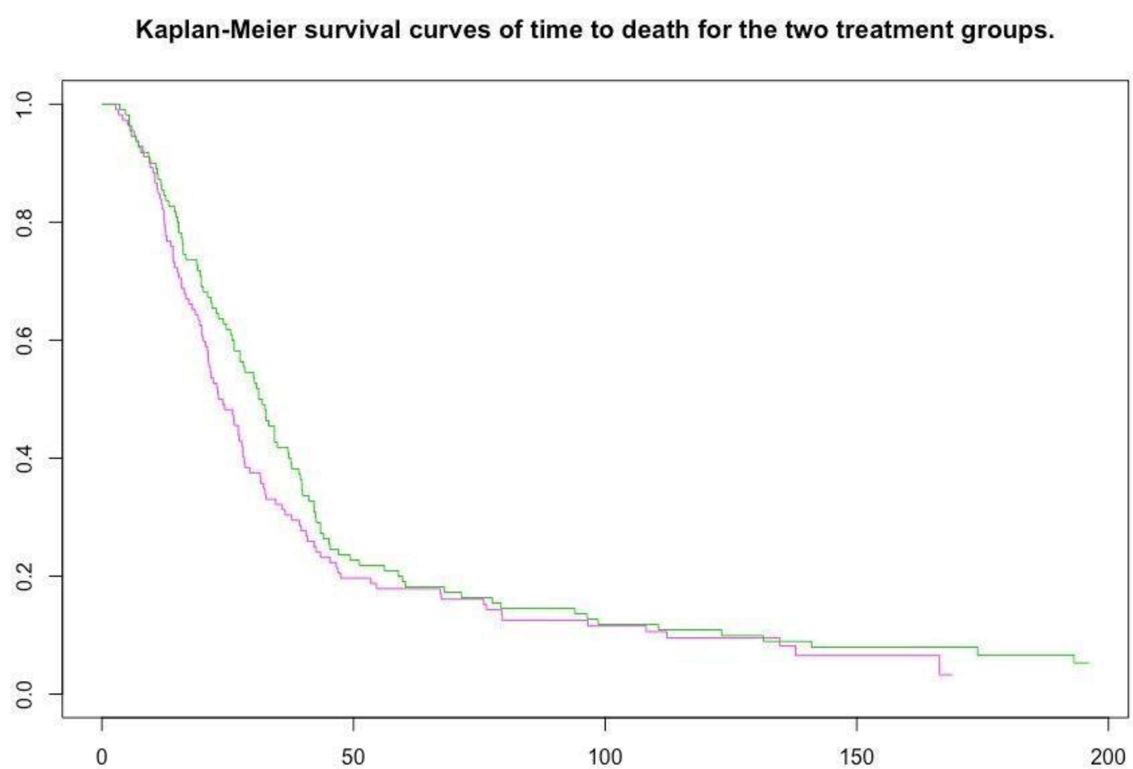
This data frame contains the following columns:

- Treat: 1= new Oeatment 0 = standard of care
- Age: age in years at the start of treatment
- Karn: 1 = Karnofsky performance score ¿ 70, 0 = Kamofsky 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

1. **Plot Kaplan-Meier survival curves of time to death for the two treatment groups.**

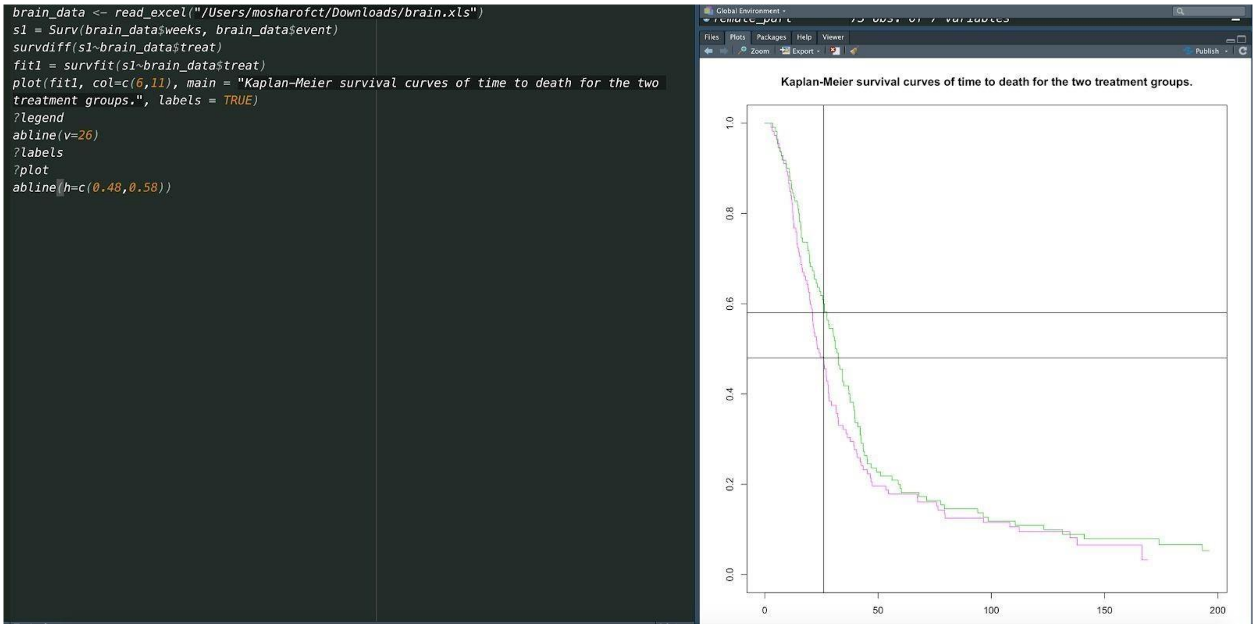
Answer:

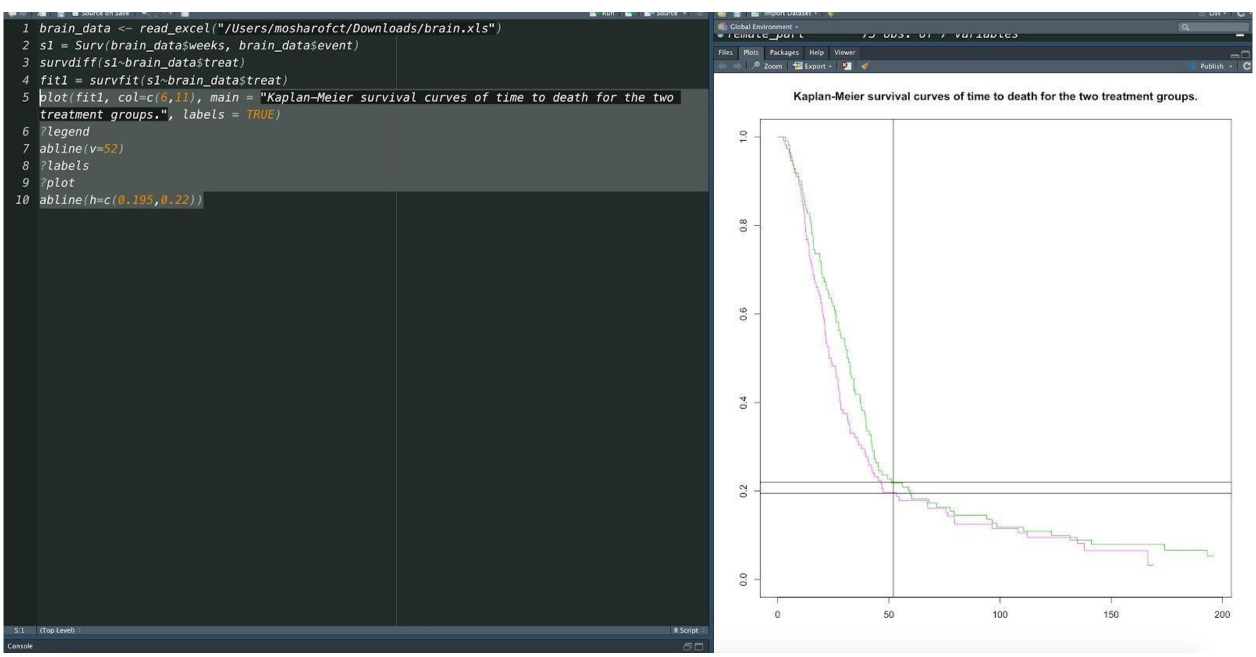
```r
brain_data <- read_excel("/Users/mosharofct/Downloads/brain.xls")
s1 = Surv(brain_data$weeks, brain_data$event)
survdiff(s1~brain_data$treat)
fit1 = survfit(s1~brain_data$treat)
plot(fit1, col=c(6,11), main = "Kaplan-Meier survival curves of time to death for the two treatment groups.")
```



1. Report the probability of death for each group at 26 weeks and 52 weeks of follow-up.

Answer:





At 26 weeks,

For old treatment, the survival probability is 0.48 And for new treatment, the survival probability is 0.58 At 52 weeks, For old treatment, the survival probability is 0.195 And for new treatment, the survival probability is 0.22

2. *Test for a difference between the two groups with the logrank test*

Answer:

```r
brain_data <- read_excel("/Users/mosharofct/Downloads/brain.xls")
s1 = Surv(brain_data$weeks, brain_data$event)
survdiff(s1~brain_data$treat)
fit1 = survfit(s1~brain_data$treat)
plot(fit1, col=c(6,11), main = "Kaplan-Meier survival curves of time to death for the two treatment groups.")
?legend
abline(v=52)
?labels
?plot
abline(h=c(0.195,0.22))
```

```
> brain_data <- read_excel("/Users/mosharofct/Downloads/brain.xls")
New names:
* `` -> ...10
> s1 = Surv(brain_data$weeks, brain_data$event)
> survdiff(s1~brain_data$treat)
Call:
survdiff(formula = s1 ~ brain_data$treat)

                     N Observed Expected (O-E)^2/E (O-E)^2/V
brain_data$treat=0 112      104     94.6     0.929      1.74
brain_data$treat=1 110      103    112.4     0.782      1.74

 Chisq= 1.7  on 1 degrees of freedom, p= 0.2
>
```

As the p-value is greater than 0.05, so, we can conclude that we don’t have enough evidence to say one treatment is better than another, no difference between the treatment.

3. Fit the semi-parametric Cox proportional hazards model of time to death, censoring individuals who were still alive at the last follow-up.

Answer:

```
brain_data <- read_excel("/Users/mosharofct/Downloads/brain.xls")
attach(brain_data)
s1 = Surv(brain_data$weeks, brain_data$event)

fit2 <- coxph(s1~factor(treat)+age+factor(karn)+factor(race)+factor(local)+factor(male)+factor(nitro)+weeks+factor(event))

summary(fit2)
```

```
> summary(fit2)
Call:
coxph(formula = s1 ~ factor(treat) + age + factor(karn) + factor(race) +
    factor(local) + factor(male) + factor(nitro) + weeks + factor(event))

  n= 221, number of events= 206
   (1 observation deleted due to missingness)

                     coef  exp(coef)   se(coef)       z Pr(>|z|)
factor(treat)1 -0.0406383  0.9601764  0.1865915  -0.218    0.828
age            -0.0007451  0.9992552  0.0080477  -0.093    0.926
factor(karn)1  -0.0108251  0.9892333  0.2047154  -0.053    0.958
factor(race)1   0.1049312  1.1106342  0.3648382   0.288    0.774
factor(local)1  0.1546114  1.1672042  0.2252017   0.687    0.492
factor(male)1  -0.0521385  0.9491974  0.1996660  -0.261    0.794
factor(nitro)1  0.1882651  1.2071535  0.2006929   0.938    0.348
weeks          -4.8729274  0.0076509  0.4239165 -11.495   <2e-16 ***
factor(event)1  3.2177668 24.9722889  5.2109421   0.618    0.537
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
                exp(coef) exp(-coef) lower .95 upper .95
factor(treat)1   0.960176    1.04148  0.666076 1.384e+00
age              0.999255    1.00075  0.983617 1.015e+00
factor(karn)1    0.989233    1.01088  0.662285 1.478e+00
factor(race)1    1.110634    0.90039  0.543276 2.271e+00
factor(local)1   1.167204    0.85675  0.750680 1.815e+00
factor(male)1    0.949197    1.05352  0.641801 1.404e+00
factor(nitro)1   1.207153    0.82840  0.814577 1.789e+00
weeks            0.007651  130.70297  0.003333 1.756e-02
factor(event)1  24.972289    0.04004  0.000916 6.808e+05

Concordance= 1  (se = 0 )
Likelihood ratio test= 1687  on 9 df,   p=<2e-16
Wald test            = 134.1  on 9 df,   p=<2e-16
Score (logrank) test = 201.8  on 9 df,   p=<2e-16
```

4. *Make a summary statement regarding the effectiveness of treatment on survival.*

Answer:

The new treatment is better effective compared to standard of care.

The independent variable weeks is statistically significant while all others are insignificant at a 5% level of significance.

The p-value for the likelihood ratio test is p=<2e-16, which is less than 0.05 implies that the model is significant at 5% level of significance.

The odds ratio for treatment is: e^(-0.0406383) = 0.96 which implies that the surviving time is better for new treatment compared to standard care.