BIOST 515 HW 7

Eliza Chai

03/08/2022

Responses

1. Figure 1 shows the Kaplan-Meier curves for the placebo and thiotepa arms with confidence intervals. We can see that the median estimated survival time is higher in thiotepa arm compared to the placebo arm. In addition, the survival curve for thiotepa arm is higher than the survival curve for the placebo arm after around 5 months. In both curves, there are censored data present in the duration of the study time. The thiotepa arm has more censored data from 30 to 60 months compare to the placebo arm.

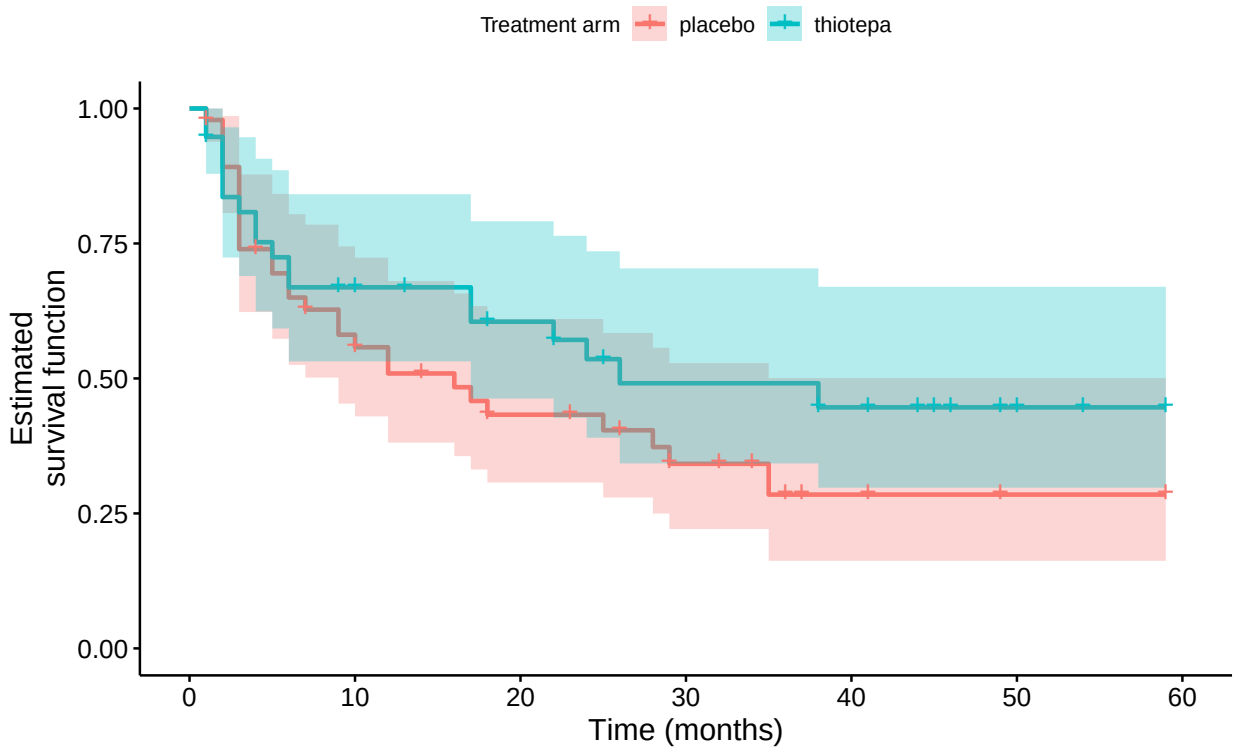

Figure 1: Kaplan-Meier curves for the placebo and thiotepa arms in the THIOTEPA dataset

2. We perform a log-rank test at a 5% significance level to compare the distribution of time to relapse in the two treatment arms: placebo arm and thiotepa arm. We fail to reject the null hypothesis that the distribution of time to relapse is equal between populations in the thiotepa arm and populations in the

placebo arm (p = 0.2). Therefore, we conclude that there is no evidence for a statistically significant difference in survival between the two treatment arms.

3. The fitted Cox proportional hazards model is as follows:

$$log(\frac{\hat{h}_i(t)}{h_0(t)}) = -0.3706 \times 1_{[treatment_i:thiotepa]}$$

$log(\frac{\hat{h}_i(t)}{h_0(t)})$ is the estimated log hazard ratio. The baseline hazard $h_0(t)$ is the hazard for a population member in the placebo arm. $h_i(t)$ is the hazard for a population member in the thiotepa arm. $1_{[treatment_i:thiotepa]} = 1$ if the participant i is in the thiotepa arm, $1_{[treatment_i:thiotepa]} = 0$ otherwise (i.e. the participant i is in the placebo arm).

The reference hazard is the hazard for a population in the placebo arm. Based on a Cox proportional hazards model, we estimate that the hazard ratio is $e^{-0.3706} = 0.690$ when comparing populations in the thiotepa arm to the placebo arm.

4. We fit a Cox proportional hazard model with log hazard ratio as the predictor and treatment as the covariate using the likelihood ratio test and performed test at a 5% significance level. Based on a Cox proportional hazards model, when comparing populations in the thiotepa treatment arm with the placebo arm (baseline hazard), we estimate that the hazard ratio is $e^{-0.3706} = 0.690$ when comparing populations in the thiotepa arm to the placebo arm (95% CI for hazard ratio with Likelihood ratio test: 0.3815, 1.2493). We fail to reject the null hypothesis that the hazard for relapse is equal between populations in the thiotepa arm and populations in the placebo arm (LRT p = 0.2153). Therefore, we conclude that there is no evidence for a statistically significant difference in hazard for relapse between the two treatment arms.

5. In (2), the null hypothesis is the distribution of time to relapse is equal between populations in the thiotepa arm and populations in the placebo arm. In (4), the hazard for relapse is equal between populations in the thiotepa arm and populations in the placebo arm. The null hypotheses tested in (2) and (4) are the same since testing if the hazard functions equal across groups is the same as testing if the survival functions are the same across populations (i.e. $h_i(t) = h_0(t)e^0 = h_0(t)$ and $h_i(t) = \frac{-\delta}{\delta t} log S_i(t)$). Hazard is the instantaneous probability of the event occurring (relapse event of first tumor recurrence) during any given time point.

6. The fitted Cox proportional hazards model is as follows:

$$log(\frac{\hat{h}_i(t)}{h_0(t)}) = -0.3677 \times 1_{[treatment_i:thiotepa]} + 0.0266 \times [tumorsize_i]$$

$log(\frac{\hat{h}_i(t)}{h_0(t)})$ is the estimated log hazard ratio. The baseline hazard $h_0(t)$ is the hazard for a population member in the placebo arm with 0 cm in size of the largest initial tumor. $h_i(t)$ is the hazard for a population member in the thiotepa arm with x cm in tumor size. $1_{[treatment_i:thiotepa]} = 1$ if the participant i is in the thiotepa arm, $1_{[treatment_i:thiotepa]} = 0$ otherwise (i.e. the participant i is in the placebo arm). $[tumorsize_i]$ is the size of largest initial tumor in centimeters.

The reference hazard is the hazard for a population in the placebo arm with 0 cm in size of the largest initial tumor. From a Cox proportional hazards model, we estimate that for two populations of the same treatment arm but differ by 1 cm in their size of largest initial tumor, the hazard ratio for relapse is $e^{0.0266} = 1.027$ when comparing the group with larger tumor size to the group with smaller tumor size (95% CI for hazard ratio with Likelihood ratio test: 0.8439, 1.2498).

Based on a Cox proportional hazards model, we estimate that for two population of the same size of largest initial tumor but in different treatment arms, the hazard ratio is $e^{-0.3677} = 0.692$ when comparing populations in the thiotepa arm to the placebo arm (95% CI for hazard ratio with Likelihood ratio test: 0.3824, 1.2536).

7. Since patients were randomly assigned to a treatment arm, this was a randomized trial. The random assignment process distributes confounding covariates among the two treatment arms equally and helps eliminate the systematic difference between groups. Therefore, the randomization would break any links between exposure and confounders and we would not be concerned about the potential confounding of the size of the largest tumor (pre-treatment).

Code Appendix

```
### Setting up the packages, options we'll need:
library(knitr)
knitr::opts_chunk$set(echo = FALSE)
library(tidyverse)
library(survival)
library(survminer)
### ------------------------------------------------------------
### Reading in the data.
thiotepa <- read_csv("~/Desktop/UW MS Biostat/BIOST Winter 2022/Biost 515/R dataset/thiotepa.csv")
### ------------------------------------------------------------
### Q1
### Create time to event variable and sort
tte <- thiotepa %>% with(Surv(stop, event))
tte %>% sort
### Plot Kaplan-Meier curves for the placebo and thiotepa arms with thiotepa dataset
survfit(tte ~ rx, thiotepa) %>%
        ggsurvplot(data = thiotepa, conf.int=T, xlab = "Time (months)",
                   ylab = "Estimated\nsurvival function",
                   legend.title = "Treatment arm",
                   legend.labs = c("placebo", "thiotepa"))
### ------------------------------------------------------------
### Q2
### Perform a log-rank test
### Null hypothesis: survival function is the same for placebo arm vs. thiotepa arm
thiotepa <- thiotepa %>% mutate(treatment = ifelse(rx == 2, 1, 0))
survdiff(tte ~ treatment, data = thiotepa)
### ------------------------------------------------------------
### Q3
### Fit a Cox proportional hazard model for THIOTEPA dataset using survival:coxph
### Treatment groups as predictor
mod1 <- coxph(tte ~ treatment, data = thiotepa)
mod1
### ------------------------------------------------------------
### Q4
### Fit a Cox proportional hazard model for treatment arm using survival:coxph
### With 95% CI LRT
mod1 %>% confint %>% exp %>% round(4)
### ------------------------------------------------------------
### Q6
### Fit a Cox proportional hazard model for THIOTEPA dataset using survival:coxph
### Treatment arms and initial tumor sizes as predictors
mod2 <- coxph(tte ~ treatment + size, data = thiotepa)
mod2
mod2 %>% confint %>% exp %>% round(4)
### ------------------------------------------------------------
```