Analysis of the factors affecting the survival time of patients with Hodgkin's lymphoma

Survival Data Analysis
2024-2025

HOMEWORK 1

2nd year Master of Statistics
Hasselt University

Group members:
Christine Katabarwa (2365938)
Mary Grace Lapurga (2365930)
Lulu Asmi Lathifah Sundayana (2365939)

Submission Date: 17 October 2024

Lecturer:
Prof. Tomasz Burzykowski

1 Introduction

This study aimed to evaluate the effect of using chemotherapy on the event-free survival time of patients with Hodgkin's lymphoma. This also investigated whether age, sex, clinical stage, mediastinum involvement, and extranodal disease have an effect on the survival time. An event in this study refers to relapse, second malignancy, or death of a patient within a 36-month period.

2 Methodology

The data of 865 Hodgkin's lymphoma patients was used in this study. Survival probabilities were estimated using Kaplan-Meier method. The confidence interval of the survival function at a given time t was estimated using the complementary log-log transformation to obtain a value within the range 0 and 1[1].

Log-rank tests were employed to assess whether survival curve estimates are different between (i) types of treatment (0-radiotherapy, 1-radiotherapy + chemotherapy); (ii) sex (1-male, 0-female); (iii) presence of extranodal disease (0-no, 1-yes); and (iv) clinical stage (1-I, 2-II). Log-rank tests for trend were used to analyze the effect of ordinal variables, namely, mediastinum involvement (0-none, 1-small, 2-large) and age groups (<26, 26-37, 37+). The original age data was recoded into three categories based on quantiles to ensure equal and sufficient sample size for the three groups for later comparison. Stratified log-rank tests were also performed to compare the survival times of patients receiving different treatments, taking into account other variables recorded for each patient. This analysis intends to remove any confounding effect that may influence the time to event.

The null hypothesis is that there is no difference in the survival time of the patients between the groups being compared. The alternative hypothesis is that there is a difference in the survival time between groups (for the log-rank tests), or there is a linear trend in the survival time over an ordinal variable (for log-rank tests for trend). All statistical tests were conducted using 5% significance level.

3 Results and Discussion

Table 1 presents the summary statistics for the overall Kaplan-Meier survival estimates for selected time periods. Standard error increases over time due to declining risk-sets, while survival probability decreases reaching 49.1% (with a 95% CI of [45.6%, 52.5%]) at month 18.

Table 1: Summary statistics for the Kaplan-Meier survival curve estimates

time	n.risk	n.event	survival	std. error	lower 95% CI	upper 95% CI
6	567	292	0.662	0.016	0.629	0.692
12	468	71	0.578	0.017	0.544	0.610
18	282	62	0.491	0.018	0.456	0.525
24	100	42	0.386	0.020	0.347	0.426
30	13	14	0.283	0.031	0.223	0.346
35	1	0	0.283	0.031	0.223	0.346

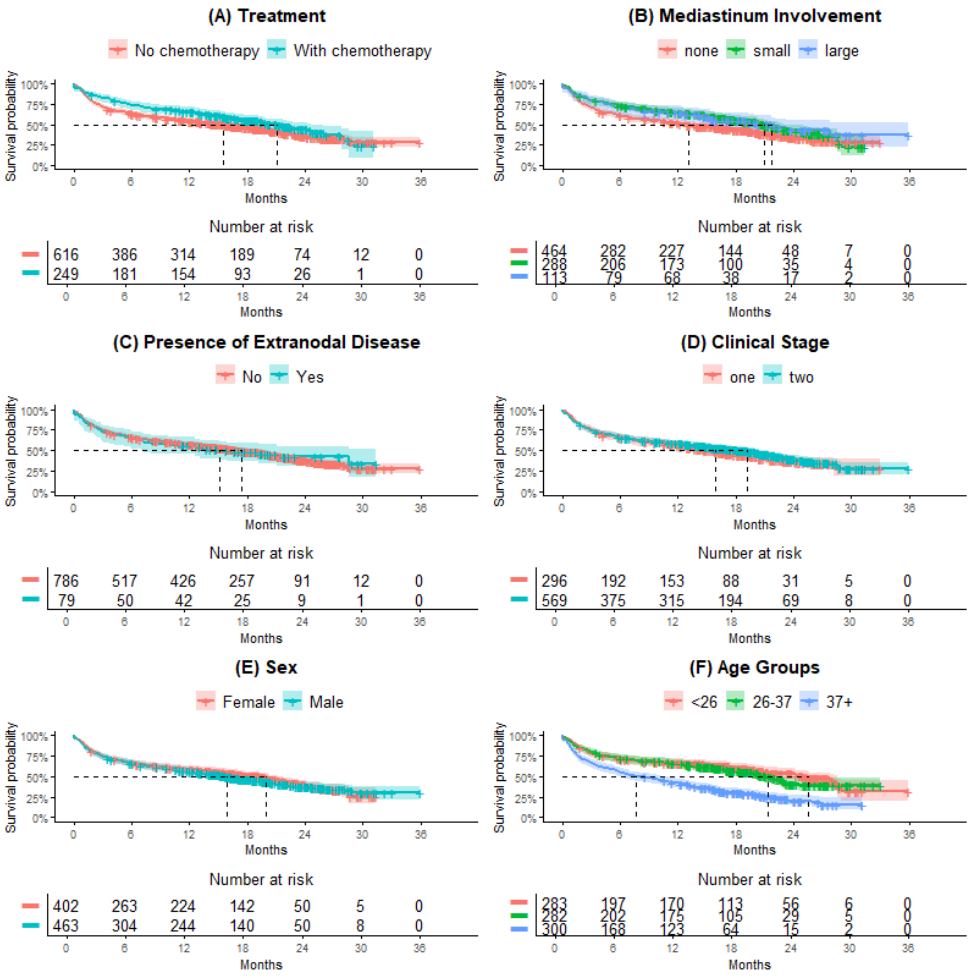

Figure 1: Kaplan-Meier Survival Curve Estimates

Figure 1 shows the plots of Kaplan-Meier survival curve estimates considering the six variables of interest. The most apparent difference in the survival curves can be seen across age groups, where older patients have lower event-free survival time than younger patients

(Figure 1F). Survival curves are lower for patients treated with only radiotherapy than those treated with both radiotherapy and chemotherapy (Figure 1A). A slight difference in survival curves can be observed among patients with small, large, or no mediastinum involvement (Figure 1B). Meanwhile, survival curve estimates seem not to differ for patients with and without extranodal disease (Figure 1C), and between stage 1 and stage 2 patients (Figure 1D). The survival estimates also seem to be the same between males and females (Figure 1E).

Table 2 summarizes the log-rank test results for each explanatory variable. The result shows that the survival time of the patients treated with chemotherapy (in the addition to radiotherapy) is significantly different from those treated with only radiotherapy. Among the level of mediastinum involvement and the level of age group, there is also a significant difference in the survival time, with a linear trend across the levels. However, there is no significant difference between the survivor functions across sex, presence of extranodal disease, and clinical stage.

Table 2: Log-rank Test Result

Variable	Treatment	Mediastinum Involvement[a]	Extranodal Disease	Clinical Stage	Sex	Age Group[a]
Test statistic	6.052	-48.092	0.002	0.846	0.575	128.314
p-value	0.014*	0.002*	0.965	0.358	0.448	<.001*

*significant at 5% level; [a]log-rank test for trend

Table 3 presents the results of the log-rank test for treatment effect, stratified by the other variables. Results revealed that there is a significant difference in the survival times of patients treated using only radiotherapy and those using both chemotherapy and radiotherapy, after allowing for the differences in the patient's sex, age, presence of extranodal disease, and clinical stage. However, there is no significant difference in the survival times of patients using either treatment after allowing for the different levels of mediastinum involvement. The insignificance of treatment difference after stratifying by mediastinum involvement could be an indication of confounding, or a result of having a smaller sample size due to stratification leading to lower power of a test.

Table 3: Log-rank Test Result for Treatment, Stratified by Different Variables

Strata	Mediastinum Involvement	Extranodal Disease	Clinical Stage	Sex	Age Group
Test statistic	3.064	6.398	5.042	6.266	5.865
p-value	0.080	0.011*	0.025*	0.012*	0.015*

*significant at 5% level

4 Conclusion

The event-free survival time of Hodgkin's lymphoma patients who are treated with radiation and chemotherapy significantly differs from those who are only treated with radiotherapy. The survival time also varies across patients with different levels of mediastinum involvement and age groups. Meanwhile, there was no sufficient evidence to conclude that survival times are different between males and females; between stage 1 and stage 2 patients; and between patients with and without extranodal disease.

References

[1] Collett, D. (2003). Modeling Survival Data in Medical Research. *Chapman and Hall/CRC*.

Appendix

```
#Sample R codes for survival plots
surv_trt <- survfit(Surv(firsttime,event)~cmt,data=hodgkindf,
type=c("kaplan-meier"), error=c("greenwood"), conf.type="log-log")
plot1 <- ggsurvplot(surv_trt,title="(A) Treatment", legend.labs =
c("No chemotherapy","With chemotherapy"), risk.table = TRUE,
surv.median.line = "hv", surv.scale="percent", break.x.by=6, conf.int=T)

#Sample SAS code for log-rank test
proc lifetest data= datasets.hodgkin plots=(s);
time firsttime*event(0); strata cmt / test=(logrank); run;

#Sample SAS code for log-rank test for trend
proc lifetest data= datasets.hodgkin plots=(s);
time firsttime*event(0); strata age_group / test=(logrank) trend; run;

#stratified logrank tests
logrank_strata1<-survdiff(Surv(firsttime, event)~cmt+strata(age_new), data=data)
logrank_strata2<-survdiff(Surv(firsttime, event)~cmt+strata(male), data=data)
logrank_strata3<-survdiff(Surv(firsttime, event)~cmt+strata(mediast), data=data)
logrank_strata4<-survdiff(Surv(firsttime, event)~cmt+strata(nodes), data=data)
logrank_strata5<-survdiff(Surv(firsttime, event)~cmt+strata(clinstg), data=data)
```