KU LEUVEN

FACULTY OF SCIENCE

Exam

Statistical Analysis of Reliability and Survival data

Daniel Gerardo GIL SANCHEZ
daniel.gilsanchez@student.kuleuven.be

Prof. Roel Braekers

Academic year 2017-2018

Questions

1. i. Give a graphical representation of the survival function for the time until infection in a person. By which time has 40% of the patients had an infection? Give the probability of not having an infection after 20 days with a 95% confidence interval (log-log-transformation)? Is this value significantly different from 75%?

The Kaplan-Meier estimator (KM) is used to give a graphical representation of the survival function.

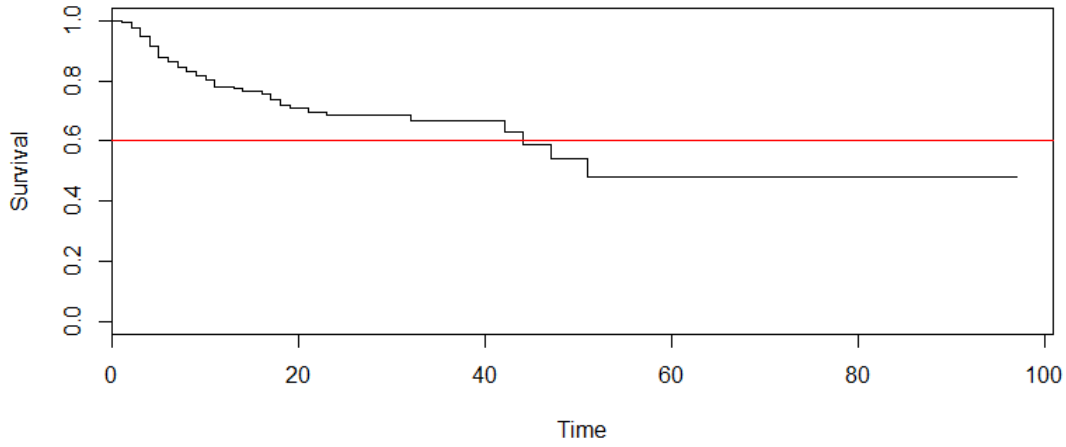


Figure 1: KM estimation of the survival function

This survival function gives the probability that an infection has not occurred by time t. To be able to know at what time 40% of the patients had a staphylococcus infection, it is enough to see at which point 60% of the patients do not have an infection, represented as a red horizontal line in Figure 1. Hence, it can be concluded that at day 44, 40% of the patients had an infection.

In a similar way, the probability of not having an infection after 20 days can be seen in a plot. Figure 2 shows the same KM survival function with log-log confidence intervals. At time 20, represented as a blue line, the probability of not having an infection is 0.71 and its confidence interval is $[0.624, 0.778]$. Note that 0.75 is inside this confidence interval, suggesting that 0.71 is not significantly different from this value.

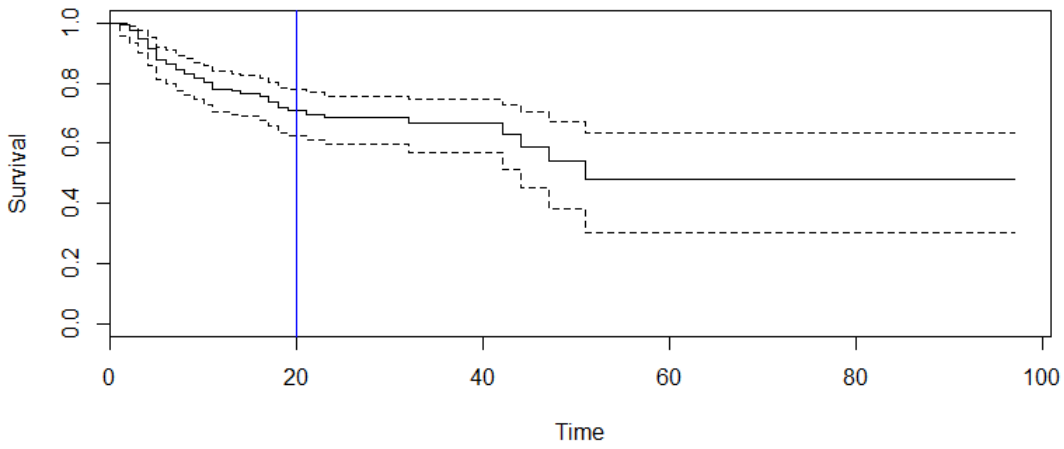


Figure 2: KM estimation of the survival function with log-log confidence interval

1. ii. Does the treatment have a significant effect on the time until infection? Show this both graphically and numerically? Based on the graph, which test-statistic should we use and why? Did you expect the numerical conclusion based on the plot?

To judge whether *treatment* produces differences on the time until the infection, the KM estimation for each group is represented in Figure 3. It shows that in the beginning of the study, the time until the infection is very similar for both treatments, actually the survival function crosses in the first five days. However, after ten days, the survival function of the patients that were treated with a body cleansing method, is above the survival function of the other group and the difference tends to be higher as time goes by.

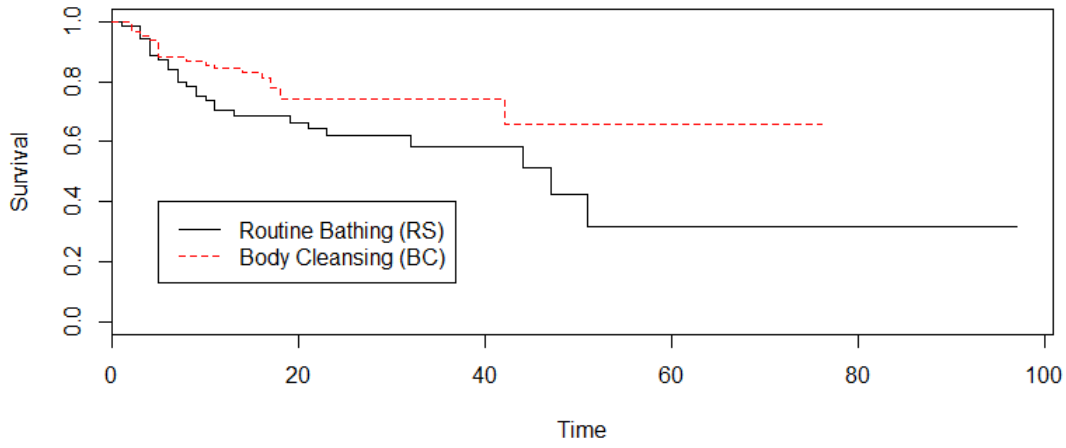


Figure 3: KM estimation of the survival function for each treatment

Until now, it is not possible to know whether this difference is statistically significant. To test this difference, the null hypothesis considered is:

$$H_0 : S_{RB}(t) = S_{BC}(t), \quad 0 < t < \tau$$

Where S_{RB} and S_{BC} are the survival functions for Routine Bathing treatment (RB) and Body Cleansing treatment (BC), respectively. And τ is the representation of the differ-

ent death times.

This hypothesis can be tested with the Tarone-Ware class of tests that under the null hypothesis follows a χ^2 distribution with one degree of freedom. It has the following structure:

$$T = \frac{\left[\sum_{l=1}^{k} w_l (d_{BC,l} - \frac{n_{BC,l} d_l}{n_l})\right]^2}{\sum_{l=1}^{k} w_l^2 \frac{n_{BC,l} n_{RS,l} d_l (n_l - d_l)}{n_l^2 (n_l - 1)}}$$

where d_l is the number of deaths at time l, $d_{BC,l}$ is the number of deaths in the body cleansing treatment at time l, n_l is the number of patients at risk at time l, $n_{BC,l}$ and $n_{RS,l}$ is the number of patients at risk in each treatment at time l and w_l are the weights for each death time.

These weights can be chosen under different criteria. For instance, the Log-rank test is a special case when all weights are set to one, i.e., all times are equally important. This scenario is optimal to detect differences in which the hazards are proportional. In contrast, the Harrington-Fleming test uses different weights to be more sensitive to late differences.

As it was described before, the difference between both treatments appears after day ten and this difference is higher over time. Besides, the fact that the survival functions cross to each other in the beginning of the study suggests that the assumption of proportionality may not hold. For these reasons the Harrington-Fleming test is used. As a result, the test is equal to 3.4 and in comparison with the χ_1^2 distribution the p-value is 0.066, suggesting that there is not difference between both treatments.

This is an unexpected result because over time the survival function of the body cleansing treatment is above the other survival function and the KM estimation shows large differences, about 30% at later times. The fact that the null hypothesis is not rejected, may lead to conclude that the confidence interval in both treatments is wider as time goes by; this can happen because the number of patients decreases over time. In appendix 2 a plot with the confidence interval for each group is shown.

1. iii. When you look at each burn site separately, does the treatment have a significant effect on the time until infection? Is gender a confounder for the influence of treatment on the time until infection?

The information on what part of the body is burnt is available, so it is of interest to see whether these *treatments* produce differences on time until infection in these parts. It is also of interest to add *gender* as confounding variable to control these results.

This analysis can be done through stratified test, which is just an extension of the tests mentioned in the previous question. Adding gender as confounding variable changes the null hypothesis because the difference between treatments is evaluated for male and female separately. This hypothesis is as follows:

$$H_0 : S_{RB,m}(t) = S_{BC,m}(t), \quad 0 < t < \tau, \quad m \in \{male, female\}$$

Using the same notation as the previous question.

From a practical point of view, separate datasets are created by filtering the patients that had burns in the corresponding part of the body. Since there are records of burns in six different parts, six separate dataset are analyzed. Results are displayed in Table 1.

Body-part	Treatment		Treatment and Gender	
	Test	p-value	Test	p-value
Head	3.12	0.08	2.23	0.14
Buttock	1.96	0.16	1.79	0.18
Trunk	2.55	0.11	3.02	0.08
Upper leg	0.71	0.4	0.95	0.33
Lower leg	3.37	0.07	4.64	0.03
Respiratory tract	5.3	0.02	6.17	0.01

Table 1: Results for each part of the body

When only *treatment* is considered, the results suggest that body cleansing is only significantly better than routine bathing on burns in the respiratory tract. For burns in other parts of the body, differences are found in the survival functions in favor of body cleansing but these are not significant.

On the other hand, when *gender* is considered as a confounding variable, statistically significant differences in *treatment* are found in burns in trunk, lower leg and respiratory tract. These results suggest that *gender* is indeed a confounding variable for some parts of the body.

In appendix 2, the plot of the estimation of the survival function for each part of the body is presented.

2. i. Is there a significant difference in time until infection between the two treatment groups? Has the fact that the time until infection is recorded in days and hence a limited amount of values for this time are seen, an influence on the method for handling the ties in the analysis?

The fact that there are more variables that may help in the explanation of whether body cleansing is a better method or not, a Cox's regression model, also known as proportional hazard model, is of interest. The variables *gender*, *race*, *percentage of the total surface burnt* and each of the *burn sites* are added to the model. In this sense, the conditional hazard is defined as:

$$\lambda(t|X) = \lambda_0(t)\exp\{\beta_1 treatment + \beta_2 gender + \beta_3 race + \beta_4 t.surface + \beta_5 head + \beta_6 buttock + \beta_7 trunk + \beta_8 upper.leg + \beta_9 lower.leg + \beta_{10} resp.tract\}$$

where $\lambda_0(t)$ is the baseline hazard.

The estimation of this model involves the computation of a partial likelihood. In this case, it is considered a partial likelihood because of the presence of censored patients, i.e., only the uncensored patients contribute in the calculation of the likelihood.

Now, this likelihood is not easy to compute when there are tied death times. For this reason, two approximations can be considered, namely Breslow and Efron. The latter is considered to be more accurate if the number of ties is large, so this approximation is used.

As a result, the Likelihood Ratio test (LR) used to check the overall significance of the model[1], gives a p-value of 0.03, suggesting that the model is significant. Furthermore, to test whether there are differences between treatments, it is sufficient to check whether β_1 is significantly different from zero (see Table 2). In this case, the p-value associated to this hypothesis is 0.03, indicating that there is a significant difference on the time until infection between both treatments.

Variable	Coefficient	Exp(coefficient)	p-value
Treatment (BC)	-0.67	0.51	0.03
Gender (female)	-0.66	0.52	0.09
Race (white)	2.02	7.56	0.05
T.surface	0.004	1.004	0.67
Head	-0.09	0.91	0.8
Buttock	0.55	1.73	0.19
Trunk	0.21	1.24	0.67
Upper.leg	-0.41	0.66	0.28
Lower.leg	-0.19	0.83	0.61
Resp.tract	0.03	1.03	0.93

Table 2: Estimation of the model

Note that the lack of precision in the measuring process of the time until infection, leads to a limited number of values of the response variable. This condition increases the number of ties present in the data, resulting in a denominator in the partial likelihood impossible to calculate because the numerical approximation may not converge. When this is the case, Efron and Breslow approximations give an idea of the real likelihood. It is well known, that the estimation from both methods are similar when there is a small number of ties, but differences appear as the number of ties increases. So the precision in how the response variable is measured has an influence on the method chosen to handle the ties.

2. ii. What is the hazard ratio of the group which burned their buttock in comparison with the non-burned group? Is it greater than one or smaller than one? What does that mean in words?

[1]The LR can be considered as the analogue of a F-test in Ordinary Least Squares.

By definition, the exponential of the coefficients of the Cox's regression model represents the proportional change in the hazard ratio with respect to the baseline hazard, when the other covariates are held constant. In this case, the exponential of the coefficient of the $buttock$ variable is 1.73 (see Table 2), a value greater than one, which indicates that a burn in a buttock increases the hazard of developing an infection 1.73 times with respect to a patient who is not burned in that part of the body. Note that even though the difference is almost double, it is not significant.

2. iii. In which direction is the influence of the other significant covariates in this model ($\alpha = 0.05$)?

Table 2 shows that only *treatment* and *race* are significant in the model. Here, the treatment of reference is routine bathing and the race of reference is non-white. The estimation of the coefficients leads to conclude that the hazard of an infection **decreases** when body cleansing treatment is used instead of the routine bathing, in fact the hazard decreases 0.51 times, when the other covariates are held constant. Regarding the race, the hazard of an infection **increases** when the patient is white. This increase is estimated as 7.56 times, when the other covariates are held constant.

3. i. Does the treatment have a significant effect on the infection time in this model? Is there an interaction between the treatment and the use of a prophylactic antibiotic treatment? Which of the other covariates has a significant effect?

The new model has the covariates *treatment*, *percentage of total surface burnt*, *time until prophylactic antibiotic treatment* and *time until excision*; in Table 3 the results are shown under the column "main effects". It can be seen that *treatment* is not significant to the model.

Variable	Main effects			With interaction		
	Coef.	Exp(coef.)	p-value	Coef.	Exp(coef.)	p-value
Treatment	-0.45	0.64	0.13	-0.25	0.78	0.62
T.surface	0.01	1.01	0.35	0.01	1.01	0.32
Time.prophylactic	0.01	1.01	0.47	0.01	1.01	0.46
Time.excision	0.01	1.01	0.32	0.01	1.01	0.26
Interaction				-0.01	0.99	0.59

Table 3: Estimation of the model

It is of interest to test whether there is an interaction between the *treatment* and *the use of a prophylactic antibiotic treatment*, so a new model is fitted adding this term, see Table 3 under the column "with interaction". As a result, none of the covariates considered is significant in the model, not even the interaction proposed.

One step further is considered, it consists in compare both models via LR test ($Chi - statistic = 0.29$ and $p - value = 0.59$). As expected, this new term does not improve significantly the fit of the model.

3. ii. Is the proportional hazards assumption satisfied for the different fixed time covariates?

To test the proportionality assumption in the fixed time covariates, an artificial time-dependent covariate is used. This variable is created as $X^*(t) = Xlog(t)$, where X can be either *treatment* or *total surface area burnt*. So in order to test the assumption for *treatment*, its corresponding artificial time-dependent covariate is added to the model (both models are considered, main effects and with interaction). As a consequence, this term is not significant in either model indicating that both treatments have proportional hazards (see Table 4).

Variable	Main effects			With interaction		
	Coef.	Exp(coef.)	p-value	Coef.	Exp(coef.)	p-value
Treatment.Artif	-0.18	0.83	0.58	-0.17	0.85	0.62
T.surface.Artif	0.01	1.01	0.32	0.01	1.01	0.32

Table 4: Estimation of time dependent variables

In a similar way, the corresponding artificial time-dependent covariate of *total surface burnt* is added to the model. This term is not significant either, which leads to conclude that the *total surface area burnt* also fulfill this assumption.

As in the previous question, a comparison between both models via LR test is considered. As expected, these new terms do not improve significantly the fit of the model, so the proportionality assumption holds in both variables.

Note that these time-dependent covariates are added to the model one at a time and not simultaneously.

4. i. In this study, we are also interested in a parametric model for the time until infection. Hereby we want to model again the time until infection as a function of treatment, gender, race, percentage of the total surface burnt and the different burn sites. Use a accelerated failure time model to find a parametric proportional hazards model with a Weibull baseline function. Which covariates have a significant effect and what does this mean in words?

The same model examined in question two is now considered with a parametric baseline hazard. It is assumed that this baseline follows a Weibull distribution, so the conditional hazard of an individual is defined as:

$$\lambda(t|X) = \lambda_0(t)\exp\{\beta_1 treatment + \beta_2 gender + \beta_3 race + \beta_4 t.surface + \beta_5 head + \beta_6 buttock + \beta_7 trunk + \beta_8 upper.leg + \beta_9 lower.leg + \beta_{10} resp.tract\}$$

This model has similar appearance to that given in question two, but there is one difference in regard to $\lambda_0(t)$. Because of the assumption that this baseline follows a Weibull distribution, it can be written as $\lambda_0(t) = \alpha\lambda t^{\alpha-1}$, where λ is the scale parameter and α is the shape parameter of the Weibull distribution. So the model can be rewritten as:

$$\lambda(t|X) = \alpha\lambda t^{\alpha-1}\exp\{\beta_1 treatment + \beta_2 gender + \beta_3 race + \beta_4 t.surface + \beta_5 head + \beta_6 buttock + \beta_7 trunk + \beta_8 upper.leg + \beta_9 lower.leg + \beta_{10} resp.tract\}$$

In software packages a different notation of this model is used. Using the log-scale notation, this model can be rewritten as:

$$Log(T) = \mu + \beta_1^* treatment + \beta_2^* gender + \beta_3^* race + \beta_4^* t.surface + \beta_5^* head + \beta_6^* buttock + \beta_7^* trunk + \beta_8^* upper.leg + \beta_9^* lower.leg + \beta_{10}^* resp.tract + \sigma E$$

Where μ and σ are the intercept and the scale parameter, respectively. And E follows an extreme value distribution.
As a consequence, the LR test used to check the global fit of the model gives a p-value of 0.016, suggesting that the model is significant. In regard to the significance of the covariates, Table 5 shows the results.

Variable	Coefficient	Std. Error	p-value
Treatment	0.72	0.36	0.05
Gender:female	0.78	0.45	0.08
Race:white	-2.37	1.17	0.04
T.surface	0	0.01	0.7
Head	0.05	0.39	0.9
Buttock	-0.68	0.48	0.15
Trunk	-0.12	0.56	0.83
Upper.leg	0.5	0.42	0.24
Lower.leg	0.3	0.43	0.48
Resp.tract	0.04	0.4	0.92
Intercept	6.09	1.29	< 0.001
log(scale)	0.12	0.12	0.34

Table 5: Estimation of the model

These results are similar to the results found in the Cox's regression model. *Treatment* is considered to be significant even though it is on the borderline of the significance level. Therefore, it is concluded that there is a significant difference in the time until infection between both treatments, when the other covariates are held constant. In fact, the sign of the parameter suggests that the time until infection when using the body cleansing treatment is longer in comparison with the routing bathing.

Likewise, there is a significant difference in the time until infection among white and non-white patients, when the other covariates are held constant. The sign of this coefficient suggests that the time until infection is shorter in white patients in comparison with non-white patients.

None of the remaining covariates considered are significant in this model.

4. ii. Compare this parametric model with the Coxs regression model in question 2. Is the Weibull distribution a good choice for the distribution of the time until infection? Show at least a fit of the baseline survival function in both models.

As a consequence of the distribution of E in the previous model using the log-scale notation, a comparison between the Cox's regression model and the accelerated failure time model can be performed. Using delta-method, the equivalence between both models is done as follows:

$$\lambda = \exp\Big(-\frac{\mu}{\sigma}\Big) = \exp(-\mu \exp(log(\sigma)))$$
$$\alpha = \frac{1}{\sigma} = \exp(-log(\sigma))$$
$$\beta = -\frac{\beta^*}{\sigma} = -\beta^* \exp(-log(\sigma))$$

In Table 6 this comparison is displayed. It seems that these results are consistent because in both models, *treatment* and *gender* are the only covariates that are significant. In regard to the point estimation of the coefficients, *treatment* changes by having 0.51 in the Cox's regression model to 0.53 in the accelerated failure time model. In gender the change is larger because in the Cox's model the estimate is 7.56, whereas in the accelerated model is 8.23.

Variable	**Accelerated model**				**Cox's model**		
	β_*	β	$\exp(\beta)$	**p-value**	β	$\exp(\beta)$	**p-value**
Treatment	0.72	−0.64	0.53	0.04	−0.67	0.51	0.03
Gender:female	0.78	−0.69	0.5	0.08	−0.66	0.52	0.09
Race:white	−2.37	2.11	8.23	0.04	2.02	7.56	0.05
T.surface	0	0	1	0.7	0	1	0.67
Head	0.05	−0.04	0.96	0.9	−0.09	0.91	0.8
Buttock	−0.68	0.6	1.83	0.15	0.55	1.73	0.19
Trunk	−0.12	0.11	1.11	0.83	0.21	1.24	0.67
Upper.leg	0.5	−0.45	0.64	0.23	−0.41	0.66	0.28
Lower.leg	0.3	−0.27	0.76	0.48	−0.19	0.83	0.61
Resp.tract	0.04	−0.04	0.96	0.92	0.03	1.03	0.93

Table 6: Comparison of coefficient estimates

To complement this comparison, AIC can be calculated for each model and see which one has the lowest measure. In Cox's regression the AIC is 438.53, whereas in the Accelerated model with the Weibull distribution the AIC is 503.87. This difference is natural since the Cox's regression is a semi-parametric procedure where the baseline is computed from the data instead of assuming any specific distribution as happens in the accelerated model. Therefore, it can be concluded that the Cox's regression has a better fit. Figure 4 shows the baseline survival function in each model.

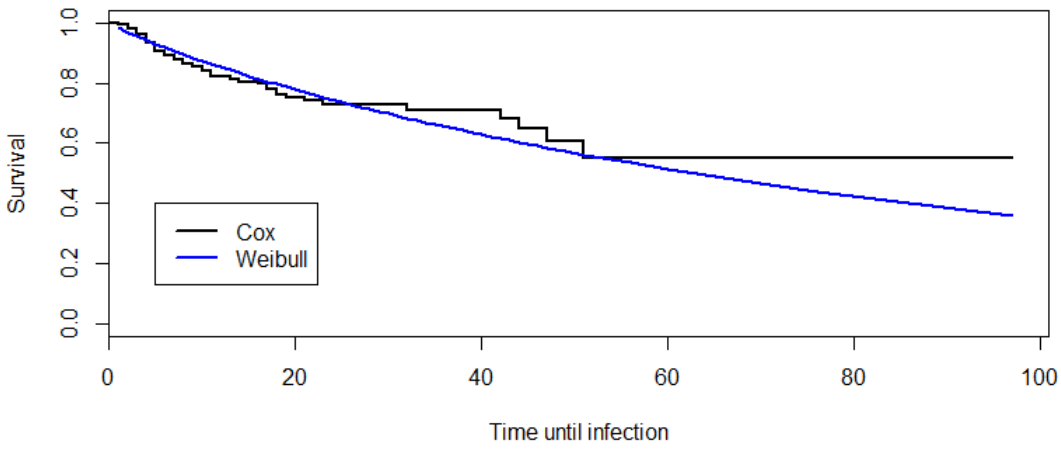


Figure 4: Comparison of baseline survival function

Despite the differences between these models, what is important here is that the body cleansing method seems to produce better results on the time until infection in comparison to the routine bathing, so it is recommended to use this procedure from now on. In addition, extra care is needed in white patients because they tend to develop the infection faster than non-white patients.

Now, in order to make sure that the Weibull distribution is a good choice for the distribution of the time until infection, a plot of the $log(-log(\hat{S}(t|X)))$ versus $log(t)$ is used (see Figure 5). It appears that the relationship between these two measures is reasonably straight, suggesting that a Weibull distribution fits the data well.

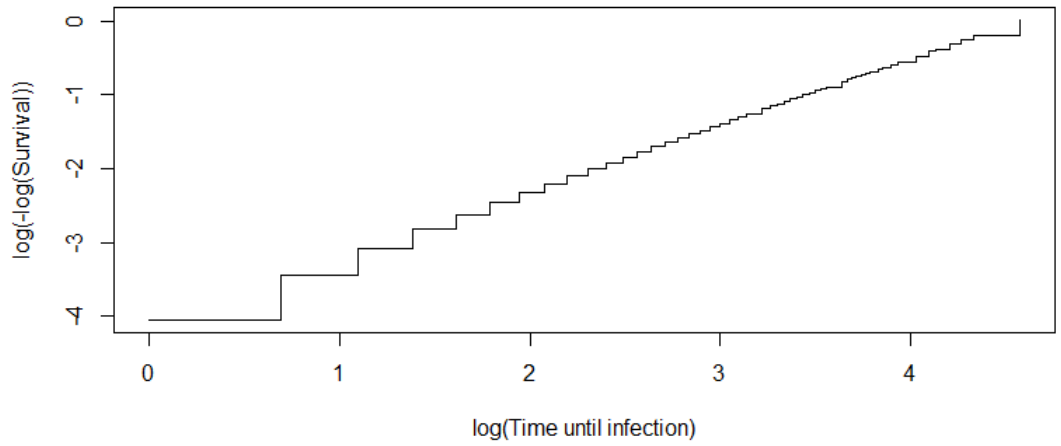


Figure 5: Goodness of fit. Weibull distribution

Appendix 1: R code

```
#   Clean workspace
rm(list=ls())

#   Set directory
setwd("...")

#   Load packages
library(survival)
library(flexsurv)

#   All comments were deleted from the txt file before importing the data
burn = read.table("BurnPatients_V2.txt",header=F)
names(burn) = c("ID","Z1","Z2","Z3","Z4","Z5","Z6","Z7","Z8","Z9","Z10","Z11",
                "T1","Ind1","T2","Ind2","T3","Ind3")

#   Descriptive statistics
#   General
summary(burn)

#   Amount of NA in each column
sapply(burn,function(x) sum(is.na(x)))

#   Amount of unique values in each column
sapply(burn, function(x) length(unique(x)))

#   Univariate analysis
#   Frequency of each variable
cbind(freq=table(burn$Z1),prop=prop.table(table(burn$Z1)))
cbind(freq=table(burn$Z2),prop=prop.table(table(burn$Z2)))
cbind(freq=table(burn$Z3),prop=prop.table(table(burn$Z3)))
cbind(freq=table(burn$Z5),prop=prop.table(table(burn$Z5)))
cbind(freq=table(burn$Z6),prop=prop.table(table(burn$Z6)))
cbind(freq=table(burn$Z7),prop=prop.table(table(burn$Z7)))
cbind(freq=table(burn$Z8),prop=prop.table(table(burn$Z8)))
cbind(freq=table(burn$Z9),prop=prop.table(table(burn$Z9)))
cbind(freq=table(burn$Z10),prop=prop.table(table(burn$Z10)))
cbind(freq=table(burn$Z11),prop=prop.table(table(burn$Z11)))
cbind(freq=table(burn$Ind1),prop=prop.table(table(burn$Ind1)))
cbind(freq=table(burn$Ind2),prop=prop.table(table(burn$Ind2)))
cbind(freq=table(burn$Ind3),prop=prop.table(table(burn$Ind3)))

#   Continuous
summary(burn[,c(5,13,15,17)])
apply(burn[,c(5,13,15,17)],2,sd)

#   1. i.

#   KM estimate
# km = survfit(Surv(burn$T3,burn$Ind3)~1,conf.type="none");km
km = survfit(Surv(burn$T3,burn$Ind3)~1);km
km_sum = summary(km);km_sum
plot(km,xlab="Time",ylab="Survival")
```

```r
#   By which time has 40% of the patients had an infection?
abline(h=0.6,col="red",lwd=1.5)
abline(v=44)

#   Give the probability of not having an infection after 20 days with a 95%
#   confidence interval (log-log-transformation)?
#   From page 73 slides
#   Log CI
tao_sq = (km_sum_20$std.err)^2/(km_sum_20$surv)^2
L = km_sum_20$surv*exp(-qnorm(0.975)*sqrt(tao_sq))
U = km_sum_20$surv*exp(qnorm(0.975)*sqrt(tao_sq))
L;U
#   log-log CI
km_ll = survfit(Surv(burn$T3,burn$Ind3)~1,conf.type="log-log");km_ll
km_ll_sum = summary(km_ll);km_ll_sum
km_ll_sum_20 = summary(km_ll,times=20);km_ll_sum_20
plot(km_ll,xlab="Time",ylab="Survival")
abline(v=20,col="blue",lwd=1.5)

# 1. ii.
#   Comparison between both treatments
km_comp = survfit(Surv(burn$T3,burn$Ind3)~burn$Z1,conf.type="none");km_comp
# km_comp = survfit(Surv(burn£T3,burn£Ind3)~burn£Z1);km_comp
km_comp_sum = summary(km_comp);km_comp_sum
plot(km_comp[1],xlab="Time",ylab="Survival")
lines(km_comp[2],col="red",lty=2)
legend(5,0.4,legend=c("Routine Bathing (RS)","Body Cleansing
 ↪  (BC)"),lty=c(1,2),col=c("black","red"))

#   Harrington-Fleming test for comparison
hf_test = survdiff(Surv(burn$T3,burn$Ind3)~burn$Z1,rho=1);hf_test

#   1. iii.
burn_Z5 = burn[burn$Z5==1,]

#   Comparison between both treatments
km_comp_Z5 =
 ↪  survfit(Surv(burn_Z5$T3,burn_Z5$Ind3)~burn_Z5$Z1,conf.type="none");km_comp_Z5
km_comp_Z5_sum = summary(km_comp_Z5);km_comp_Z5_sum
par(mfrow=c(1,1))
plot(km_comp_Z5[1],xlab="Time",ylab="Survival")
lines(km_comp_Z5[2],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),col=c("black","red"))

#   Comparison between both treatments
hf_test_Z5 = survdiff(Surv(burn_Z5$T3,burn_Z5$Ind3)~burn_Z5$Z1,rho=1);hf_test_Z5

#   Controlling for gender
km_comp_Z5_gender = survfit(Surv(burn_Z5$T3,burn_Z5$Ind3)~burn_Z5$Z1+
    strata(burn_Z5$Z2),conf.type="none");km_comp_Z5_gender
km_comp_Z5_age_sum = summary(km_comp_Z5_gender);km_comp_Z5_age_sum
par(mfrow = c(1,2))
plot(km_comp_Z5_gender[1],xlab="Time",ylab="Survival",main="Male")
lines(km_comp_Z5_gender[3],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),
```

```
       col=c("black","red"))
plot(km_comp_Z5_gender[2],xlab="Time",ylab="Survival",main="Female")
lines(km_comp_Z5_gender[4],col="red",lty=2)
legend(5,0.3,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))

#   Comparison between both treatments
hf_test_Z5_gender = survdiff(Surv(burn_Z5$T3,burn_Z5$Ind3)~burn_Z5$Z1+
                                 strata(burn_Z5$Z2),rho=1);hf_test_Z5_gender


##############
#   Z6==1: Burns in the butt
burn_Z6 = burn[burn$Z6==1,]

#   Comparison between both treatments
km_comp_Z6 =
 ↪  survfit(Surv(burn_Z6$T3,burn_Z6$Ind3)~burn_Z6$Z1,conf.type="none");km_comp_Z6
km_comp_Z6_sum = summary(km_comp_Z6);km_comp_Z6_sum
par(mfrow=c(1,1))
plot(km_comp_Z6[1],xlab="Time",ylab="Survival")
lines(km_comp_Z6[2],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),col=c("black","red"))

#   Comparison between both treatments
hf_test_Z6 = survdiff(Surv(burn_Z6$T3,burn_Z6$Ind3)~burn_Z6$Z1,rho=1);hf_test_Z6

#   Controlling for gender
km_comp_Z6_gender = survfit(Surv(burn_Z6$T3,burn_Z6$Ind3)~burn_Z6$Z1+
    strata(burn_Z6$Z2),conf.type="none");km_comp_Z6_gender
km_comp_Z6_age_sum = summary(km_comp_Z6_gender);km_comp_Z6_age_sum
par(mfrow = c(1,2))
plot(km_comp_Z6_gender[1],xlab="Time",ylab="Survival",main="Male")
lines(km_comp_Z6_gender[3],col="red",lty=2)
legend(5,0.3,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))
plot(km_comp_Z6_gender[2],xlab="Time",ylab="Survival",main="Female")
lines(km_comp_Z6_gender[4],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))

#   Comparison between both treatments
hf_test_Z6_gender = survdiff(Surv(burn_Z6$T3,burn_Z6$Ind3)~burn_Z6$Z1+
                                 strata(burn_Z6$Z2),rho=1);hf_test_Z6_gender

##############
#   Z7==1: Burns in the trunk
burn_Z7 = burn[burn$Z7==1,]

#   Comparison between both treatments
km_comp_Z7 =
 ↪  survfit(Surv(burn_Z7$T3,burn_Z7$Ind3)~burn_Z7$Z1,conf.type="none");km_comp_Z7
km_comp_Z7_sum = summary(km_comp_Z7);km_comp_Z7_sum
par(mfrow=c(1,1))
plot(km_comp_Z7[1],xlab="Time",ylab="Survival")
```

```
lines(km_comp_Z7[2],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),col=c("black","red"))

#   Comparison between both treatments
hf_test_Z7 = survdiff(Surv(burn_Z7$T3,burn_Z7$Ind3)~burn_Z7$Z1,rho=1);hf_test_Z7

#   Controlling for gender
km_comp_Z7_gender = survfit(Surv(burn_Z7$T3,burn_Z7$Ind3)~burn_Z7$Z1+
    strata(burn_Z7$Z2),conf.type="none");km_comp_Z7_gender
km_comp_Z7_age_sum = summary(km_comp_Z7_gender);km_comp_Z7_age_sum
par(mfrow = c(1,2))
plot(km_comp_Z7_gender[1],xlab="Time",ylab="Survival",main="Male")
lines(km_comp_Z7_gender[3],col="red",lty=2)
legend(2,0.4,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))
plot(km_comp_Z7_gender[2],xlab="Time",ylab="Survival",main="Female")
lines(km_comp_Z7_gender[4],col="red",lty=2)
legend(5,0.3,llegend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))

#   Comparison between both treatments
hf_test_Z7_gender = survdiff(Surv(burn_Z7$T3,burn_Z7$Ind3)~burn_Z7$Z1+
                                strata(burn_Z7$Z2),rho=1);hf_test_Z7_gender

##############
#   Z8==1: Burns in the upper leg
burn_Z8 = burn[burn$Z8==1,]

#   Comparison between both treatments
km_comp_Z8 =
 ↪  survfit(Surv(burn_Z8$T3,burn_Z8$Ind3)~burn_Z8$Z1,conf.type="none");km_comp_Z8
km_comp_Z8_sum = summary(km_comp_Z8);km_comp_Z8_sum
par(mfrow=c(1,1))
plot(km_comp_Z8[1],xlab="Time",ylab="Survival")
lines(km_comp_Z8[2],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),col=c("black","red"))

#   Comparison between both treatments
hf_test_Z8 = survdiff(Surv(burn_Z8$T3,burn_Z8$Ind3)~burn_Z8$Z1,rho=1);hf_test_Z8

#   Controlling for gender
km_comp_Z8_gender = survfit(Surv(burn_Z8$T3,burn_Z8$Ind3)~burn_Z8$Z1+
    strata(burn_Z8$Z2),conf.type="none");km_comp_Z8_gender
km_comp_Z8_age_sum = summary(km_comp_Z8_gender);km_comp_Z8_age_sum
par(mfrow = c(1,2))
plot(km_comp_Z8_gender[1],xlab="Time",ylab="Survival",main="Male")
lines(km_comp_Z8_gender[3],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))
plot(km_comp_Z8_gender[2],xlab="Time",ylab="Survival",main="Female")
lines(km_comp_Z8_gender[4],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))

#   Comparison between both treatments
```

```r
hf_test_Z8_gender = survdiff(Surv(burn_Z8$T3,burn_Z8$Ind3)~burn_Z8$Z1+
                                strata(burn_Z8$Z2),rho=1);hf_test_Z8_gender

##############
#   Z9==1: Burns in the lower leg
burn_Z9 = burn[burn$Z9==1,]

#   Comparison between both treatments
km_comp_Z9 =
 ↪  survfit(Surv(burn_Z9$T3,burn_Z9$Ind3)~burn_Z9$Z1,conf.type="none");km_comp_Z9
km_comp_Z9_sum = summary(km_comp_Z9);km_comp_Z9_sum
par(mfrow=c(1,1))
plot(km_comp_Z9[1],xlab="Time",ylab="Survival")
lines(km_comp_Z9[2],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),col=c("black","red"))

#   Comparison between both treatments
hf_test_Z9 = survdiff(Surv(burn_Z9$T3,burn_Z9$Ind3)~burn_Z9$Z1,rho=1);hf_test_Z9

#   Controlling for gender
km_comp_Z9_gender = survfit(Surv(burn_Z9$T3,burn_Z9$Ind3)~burn_Z9$Z1+
    strata(burn_Z9$Z2),conf.type="none");km_comp_Z9_gender
km_comp_Z9_age_sum = summary(km_comp_Z9_gender);km_comp_Z9_age_sum
par(mfrow = c(1,2))
plot(km_comp_Z9_gender[1],xlab="Time",ylab="Survival",main="Male")
lines(km_comp_Z9_gender[3],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))
plot(km_comp_Z9_gender[2],xlab="Time",ylab="Survival",main="Female")
lines(km_comp_Z9_gender[4],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))

#   Comparison between both treatments
hf_test_Z9_gender = survdiff(Surv(burn_Z9$T3,burn_Z9$Ind3)~burn_Z9$Z1+
                                strata(burn_Z9$Z2),rho=1);hf_test_Z9_gender

##############
#   Z10==1: Burns in the respiratory tract
burn_Z10 = burn[burn$Z10==1,]

#   Comparison between both treatments
km_comp_Z10 =
 ↪  survfit(Surv(burn_Z10$T3,burn_Z10$Ind3)~burn_Z10$Z1,conf.type="none")
km_comp_Z10
km_comp_Z10_sum = summary(km_comp_Z10);km_comp_Z10_sum
par(mfrow=c(1,1))
plot(km_comp_Z10[1],xlab="Time",ylab="Survival")
lines(km_comp_Z10[2],col="red",lty=2)
legend(5,0.3,legend=c("RB","BC"),lty=c(1,2),col=c("black","red"))

#   Comparison between both treatments
hf_test_Z10 =
 ↪  survdiff(Surv(burn_Z10$T3,burn_Z10$Ind3)~burn_Z10$Z1,rho=1);hf_test_Z10
```

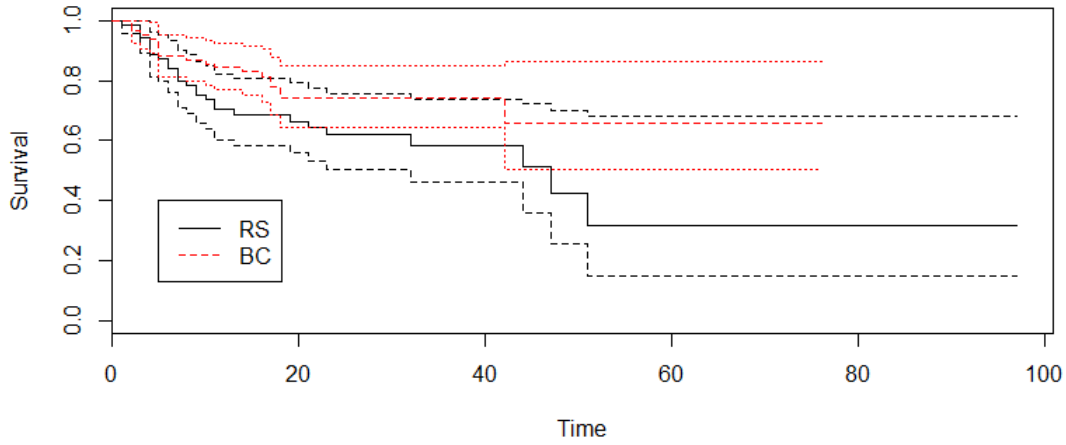
```
#   Controlling for gender
km_comp_Z10_gender = survfit(Surv(burn_Z10$T3,burn_Z10$Ind3)~burn_Z10$Z1+
    strata(burn_Z10$Z2),conf.type="none");km_comp_Z10_gender
km_comp_Z10_age_sum = summary(km_comp_Z10_gender);km_comp_Z10_age_sum
par(mfrow = c(1,2))
plot(km_comp_Z10_gender[1],xlab="Time",ylab="Survival",main="Male")
lines(km_comp_Z10_gender[3],col="red",lty=2)
legend(5,0.4,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))
plot(km_comp_Z10_gender[2],xlab="Time",ylab="Survival",main="Female")
lines(km_comp_Z10_gender[4],col="red",lty=2)
legend(5,0.23,legend=c("RB","BC"),lty=c(1,2),
       col=c("black","red"))

#   Comparison between both treatments
hf_test_Z10_gender = survdiff(Surv(burn_Z10$T3,burn_Z10$Ind3)~burn_Z10$Z1+
                                  strata(burn_Z10$Z2),rho=1);hf_test_Z10_gender


#   Combining all results
result = rbind(head = c(hf_test_Z5$chisq,pchisq(hf_test_Z5$chisq,1,lower.tail =
 ↪  F),hf_test_Z5_gender$chisq,pchisq(hf_test_Z5_gender$chisq,1,lower.tail = F)),
               Buttock = c(hf_test_Z6$chisq,pchisq(hf_test_Z6$chisq,1,lower.tail =
                ↪  F),hf_test_Z6_gender$chisq,pchisq(hf_test_Z6_gender$chisq,1,
               lower.tail = F)),
               Trunk = c(hf_test_Z7$chisq,pchisq(hf_test_Z7$chisq,1,lower.tail =
                ↪  F),hf_test_Z7_gender$chisq,pchisq(hf_test_Z7_gender$chisq,1,
               lower.tail = F)),
               Upper_leg = c(hf_test_Z8$chisq,pchisq(hf_test_Z8$chisq,1,lower.tail
                ↪  = F),hf_test_Z8_gender$chisq,pchisq(hf_test_Z8_gender$chisq,1,
               lower.tail = F)),
               Lower_leg = c(hf_test_Z9$chisq,pchisq(hf_test_Z9$chisq,1,lower.tail
                ↪  = F),hf_test_Z9_gender$chisq,pchisq(hf_test_Z9_gender$chisq,1,
               lower.tail = F)),
               Respiratory_tract =
                ↪  c(hf_test_Z10$chisq,pchisq(hf_test_Z10$chisq,1,lower.tail =
                ↪  F),hf_test_Z10_gender$chisq,pchisq(hf_test_Z10_gender$chisq,1,
               lower.tail = F)))
               colnames(result)=c("chisq","pvalue","gender chisq","gender
                ↪  pvalue");result



# 2. i.
# View(sort(Surv(burn£T3,burn£Ind3))) # There are a lot of ties
cox = coxph(Surv(burn$T3,burn$Ind3)~burn$Z1+burn$Z2+burn$Z3+burn$Z4+burn$Z5+
              burn$Z6+burn$Z7+burn$Z8+burn$Z9+burn$Z10);cox
cox_sum = summary(cox);cox_sum

# Plot the baseline survival function
plot(survfit(cox),xlab = "Time until infection",ylab = "Survival")

# cox2 = coxph(Surv(burn£T3,burn£Ind3)~burn£Z1+burn£Z3);cox2
# cox2_sum = summary(cox2);cox2_sum
# anova(cox,cox2)
```

```
#   2. ii.
cox_sum

#   2. iii.
cox_sum

#   3.i.
cox3 = coxph(Surv(burn$T3,burn$Ind3)~burn$Z1+burn$Z4+burn$T1+burn$T2);cox3
cox3_sum = summary(cox3);cox3_sum

#   Add interaction
cox4 = update(cox3,.~.+burn$Z1*burn$T2)
cox4_sum = summary(cox4);cox4_sum
anova(cox3,cox4)
#   The interaction is not significant using LR test
#   No variable is significant

#   3. ii. Is the proportional hazards assumption satisfied for the different
 ↪  fixed time covariates?
cox.zph(cox3)
cox.zph(cox4)

#   From Slide 213
#   Main effects model
cox3_z1 =
 ↪  coxph(Surv(burn$T3,burn$Ind3)~burn$Z1+burn$Z4+burn$T1+burn$T2+tt(burn$Z1),
tt=function(x,t,...)x*log(t))
summary(cox3_z1)
cox3_z4 =
 ↪  coxph(Surv(burn$T3,burn$Ind3)~burn$Z1+burn$Z4+burn$T1+burn$T2+tt(burn$Z4),
tt=function(x,t,...)x*log(t))
summary(cox3_z4)
#   The proportional hazard assumption holds for both covariates.

#   Model with interactions
cox4_z1 =
 ↪  coxph(Surv(burn$T3,burn$Ind3)~burn$Z1+burn$Z4+burn$T1+burn$T2+burn$Z1*burn$T2+
tt(burn$Z1),
tt=function(x,t,...)x*log(t))
summary(cox4_z1)
cox4_z4 =
 ↪  coxph(Surv(burn$T3,burn$Ind3)~burn$Z1+burn$Z4+burn$T1+burn$T2+burn$Z1*burn$T2+
tt(burn$Z4),
tt=function(x,t,...)x*log(t))
summary(cox4_z4)


# 4. i.
wei = survreg(Surv(burn$T3,burn$Ind3)~burn$Z1+burn$Z2+burn$Z3+burn$Z4+burn$Z5+
                  burn$Z6+burn$Z7+burn$Z8+burn$Z9+burn$Z10);wei
wei_sum = summary(wei);wei_sum

#  4. ii.
#   Comparison of estimates
```

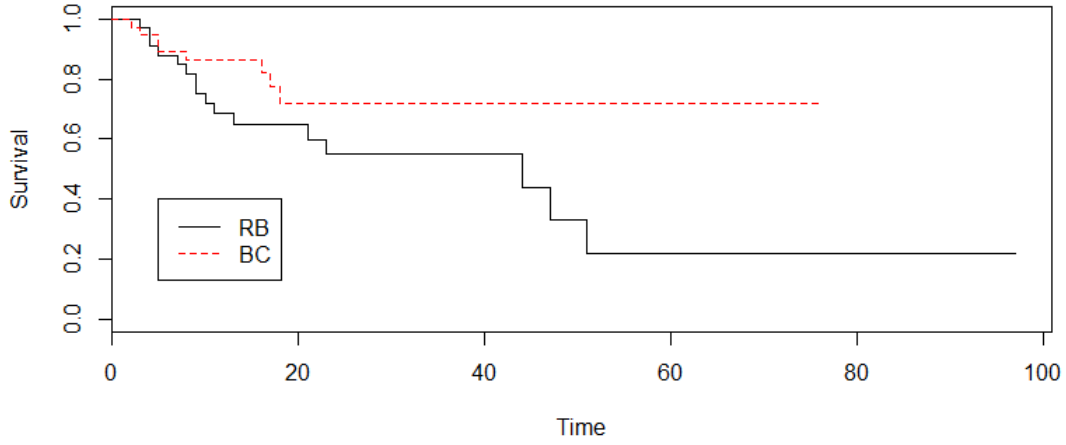
```
para = wei$coef
lscale = log(wei$scale)
V = wei$var
lambda = exp(-para[1]*exp(-lscale))
alpha = exp(-lscale)
beta = -para[-1]*exp(-lscale)
x<-c(lambda,alpha,beta)
names(x)[1]<-"lambda"
names(x)[2]<-"alpha"
x

m = length(para[-1]) #  Number of covariates
G = matrix(0,nrow=m+2,ncol=m+2)
G[1,1] = -exp(-para[1]*exp(-lscale))*exp(-lscale)
G[2:(m+1),3:(m+2)] = diag(m)*(-exp(-lscale))
G[m+2,1] = exp(-para[1]*exp(-lscale))*para[1]*exp(-lscale)
G[m+2,2] = -exp(-lscale)
G[m+2,3:(m+2)] = para[-1]*exp(-lscale)
G

PrVar = t(G)%*%V%*%G;PrVar
PrStd = sqrt(diag(PrVar));PrStd
PrChisq = c(" "," ",(x[3:(m+2)]/PrStd[3:(m+2)])^2)
PrPvalue = c(" "," ",pchisq((x[3:(m+2)]/PrStd[3:(m+2)])^2,1,lower.tail=F))
out = data.frame(x,PrStd,PrChisq,PrPvalue)
names(out) = c("Estimate","StdError","Chisq","P-value")
out
out$exp_estimate = exp(out$Estimate)

#   AIC comparison
extractAIC(cox)
extractAIC(wei)

#   Baseline in each model
par(mfrow=c(1,1))
# Cox model
plot(survfit(cox,type="breslow",conf.type = "none"),xlab = "Time until
 ↪  infection",ylab = "Survival",lwd=2)
# Weibull model
wei_2 =
 ↪  flexsurvreg(Surv(burn$T3,burn$Ind3)~burn$Z1+burn$Z2+burn$Z3+burn$Z4+burn$Z5+
      burn$Z6+burn$Z7+burn$Z8+burn$Z9+burn$Z10,dist='weibull')
wei_2_sum = summary(wei_2,tidy=T)
lines(wei_2_sum$time,wei_2_sum$est,col="blue",lwd=2)
legend(5,0.4,legend=c("Cox","Weibull"),lwd=c(2,2),col=c("black","blue"))

#   Is the Weibull distribution a good choice for the distribution of the time
 ↪  until infection?
#   From page 290
x = as.matrix(burn[,2:11],ncol=10)
coeff = matrix(para[2:11],ncol=10)
beta_x = x%*%t(coeff)

surv_est = exp(-lambda*(burn$T3^(alpha))*exp((beta_x)))
plot(log(burn$T3), log(-log(surv_est)), col='black', xlab = 'log(Time until
 ↪  infection)',ylab = 'log(-log(Survival Curve))')
```

Appendix 2: Computer - output

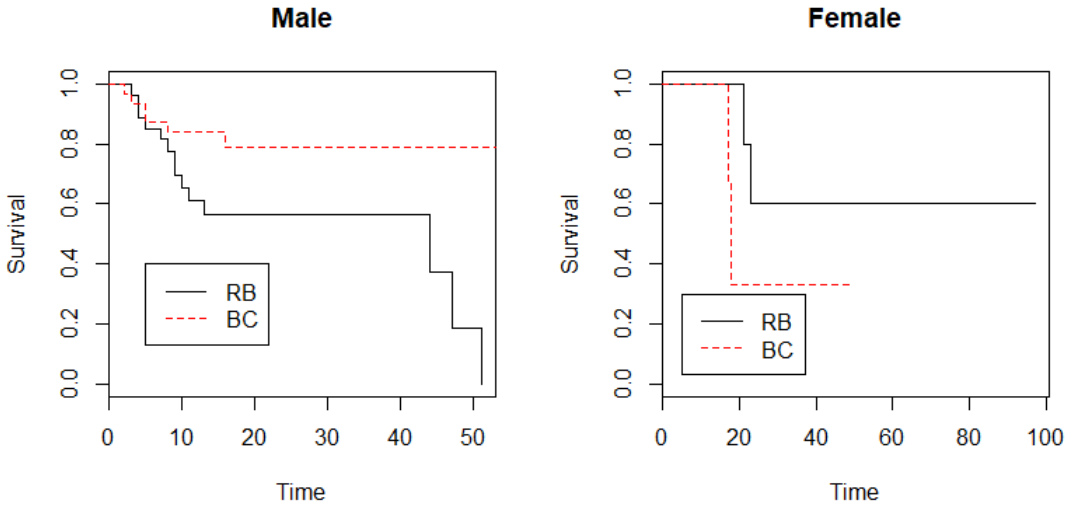
- Comparison between treatments with confidence intervals



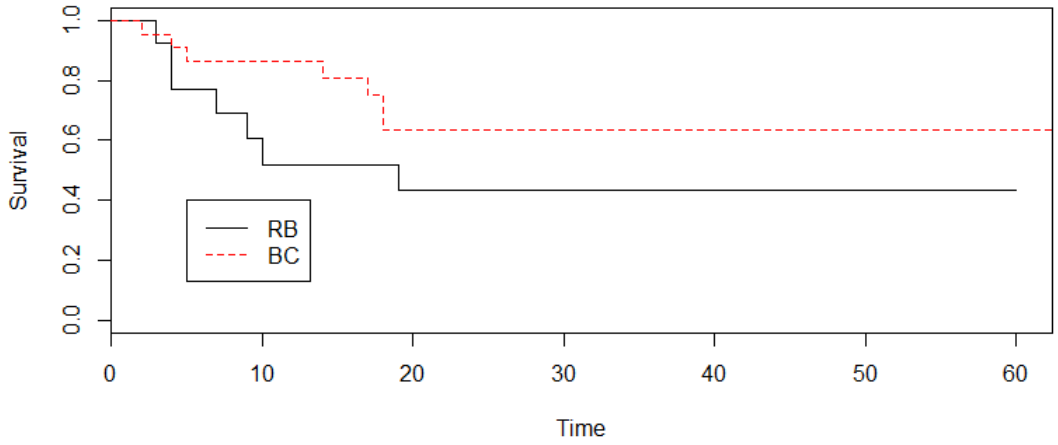
- Survival Function burns in the Head



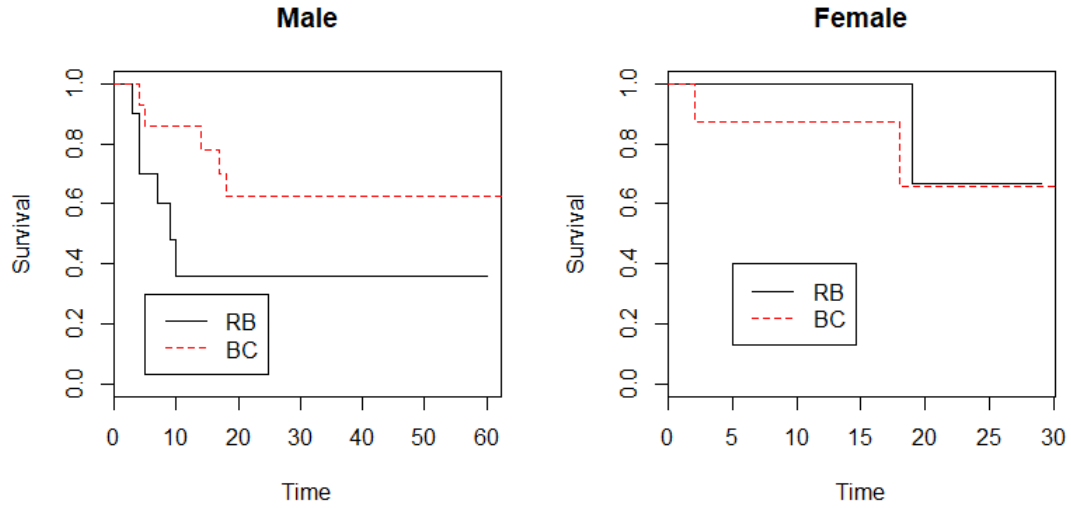
- Survival Function burns in the Head by Gender



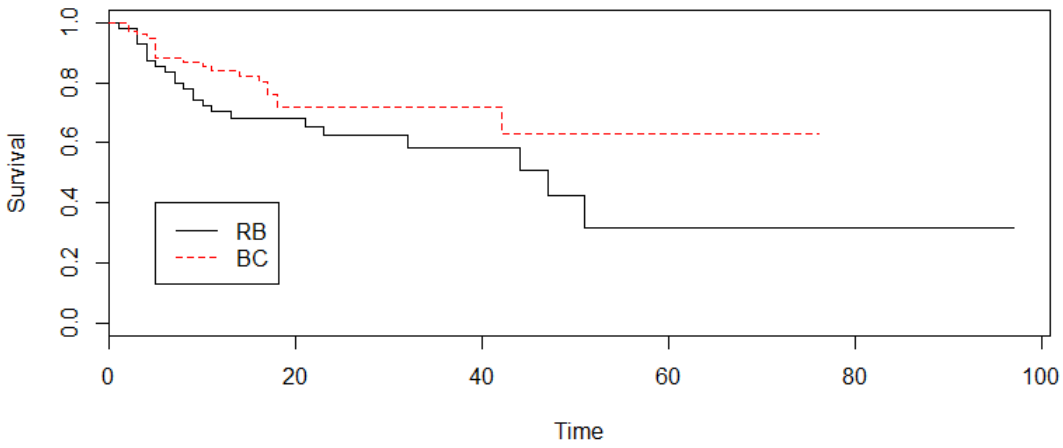
- Survival Function burns in the Buttock



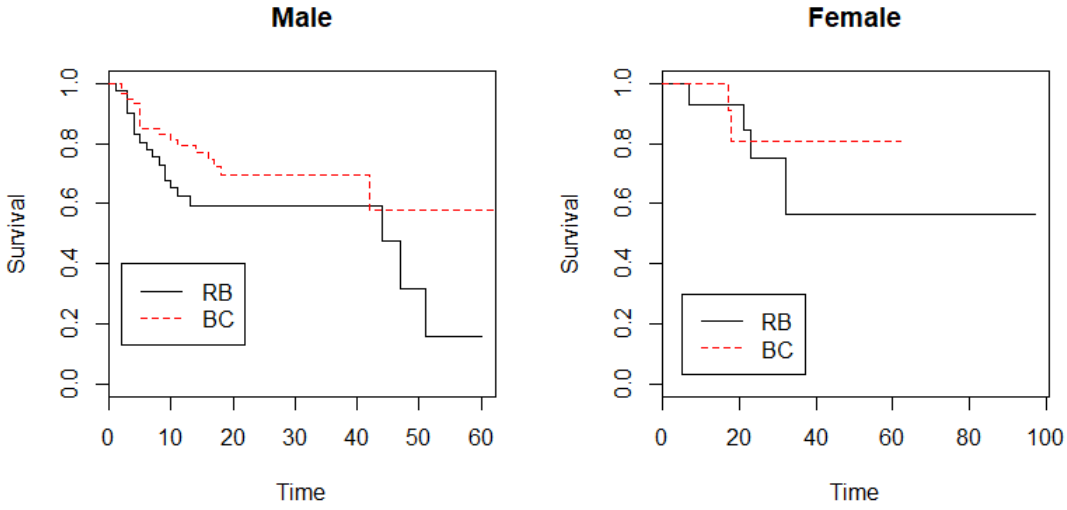
- Survival Function burns in the Buttock by Gender



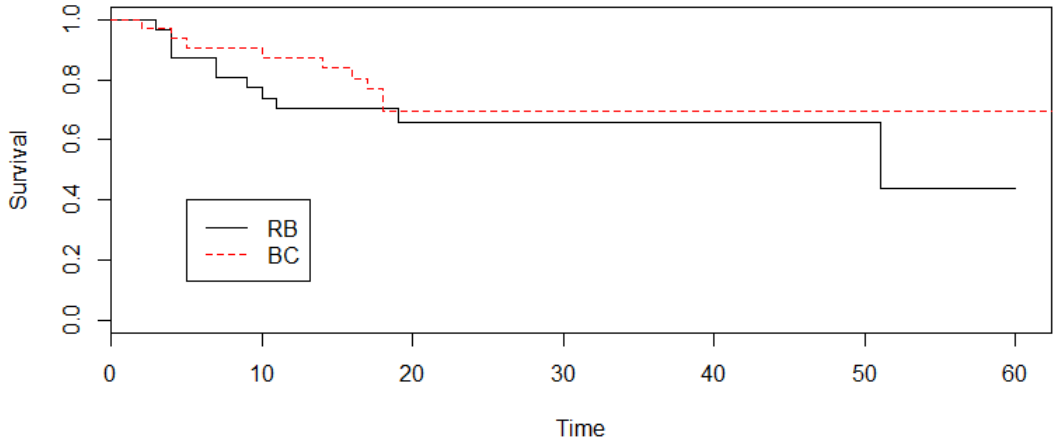
- Survival Function burns in the Trunk



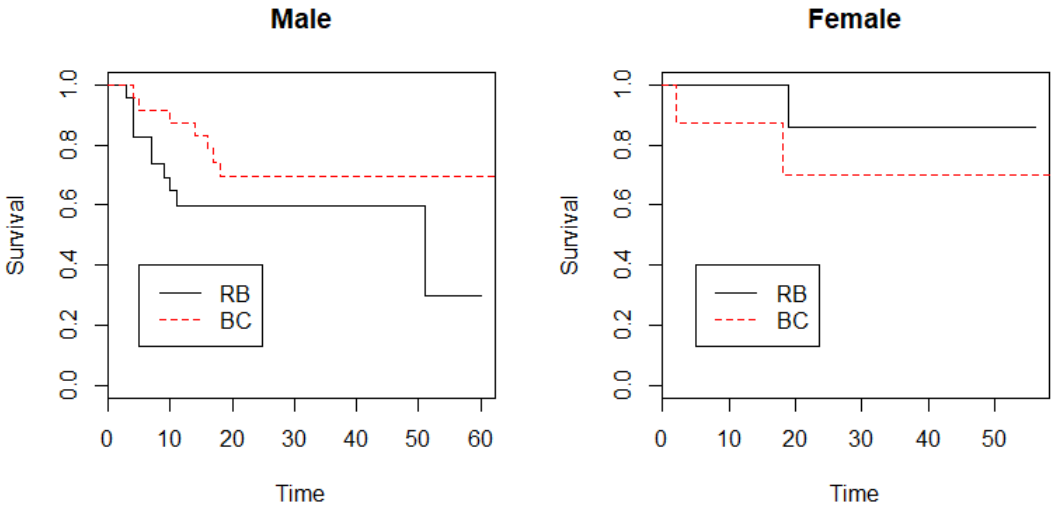
- Survival Function burns in the Trunk by Gender



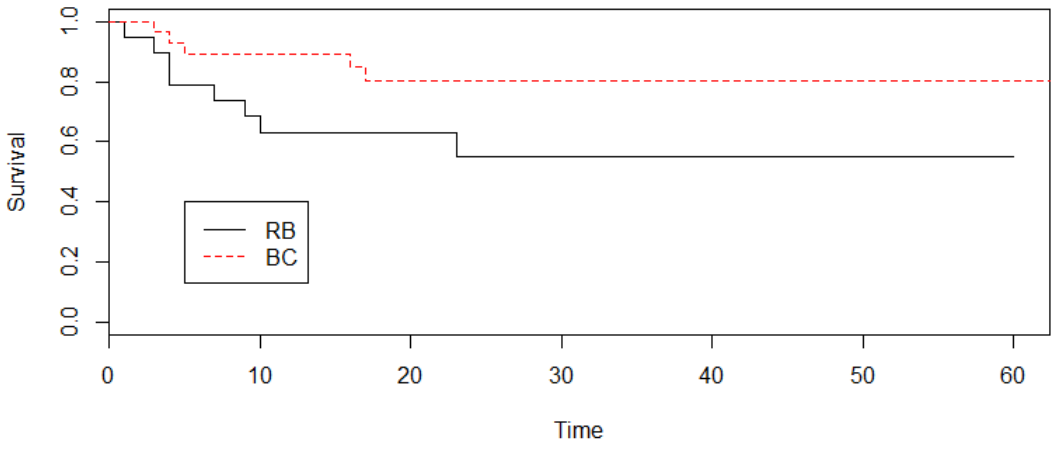
- Survival Function burns in the Upper leg



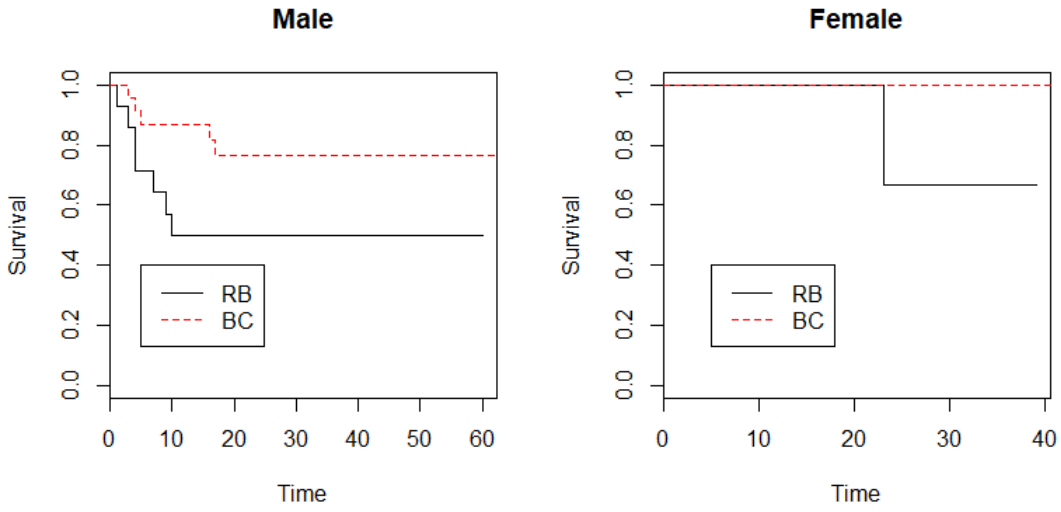
- Survival Function burns in the Upper leg by Gender



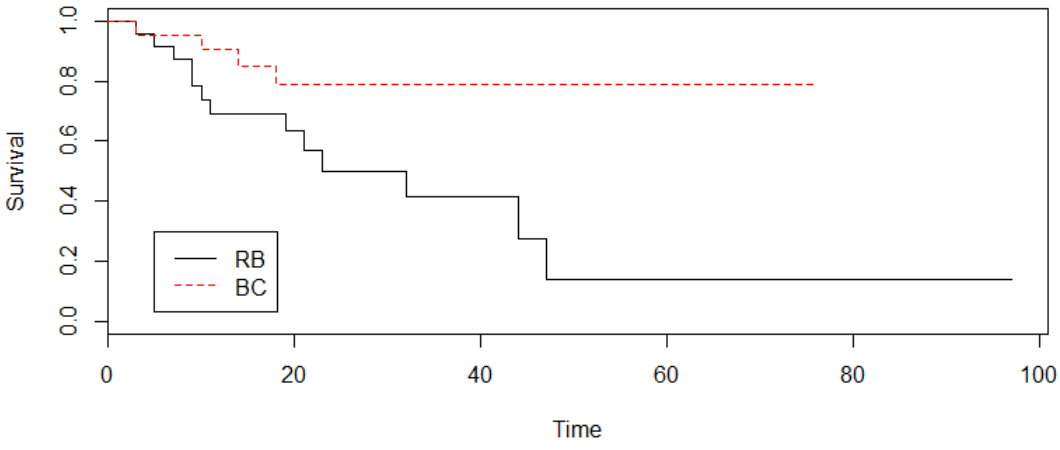
- Survival Function burns in the Lower leg



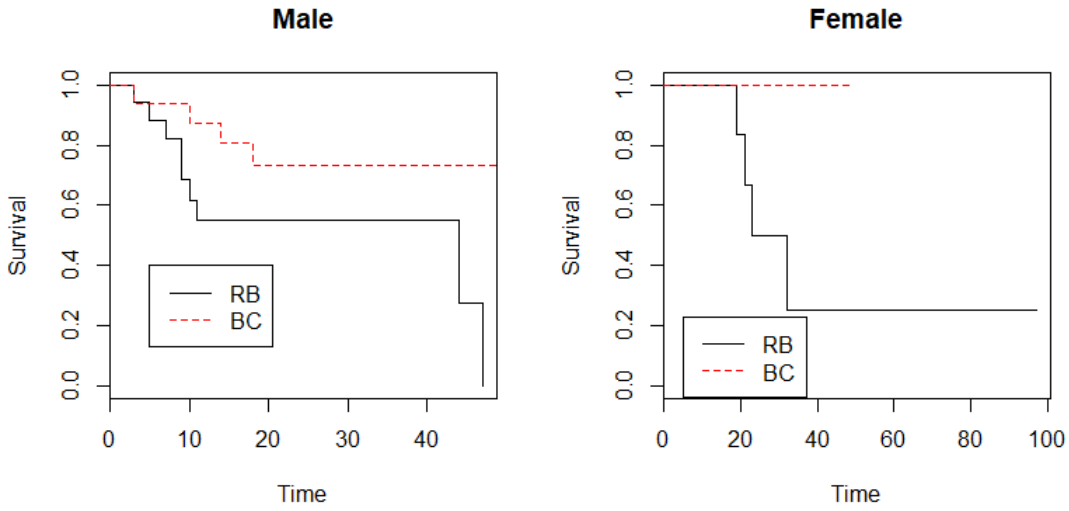
- Survival Function burns in the Lower leg by Gender



- Survival Function burns in the Respiratory Tract



- Survival Function burns in the Respiratory Tract by Gender


Male
Survival
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
Time
RB
BC
Female
Survival
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
Time
RB
BC