Paper CT09

One sample Log-Rank test for one sample survival studies: SAS code vs R code

Giulio Pacini, Menarini Ricerche, Florence, Italy
Giulia Tonini, Menarini Ricerche, Florence, Italy
Simona Scartoni, Menarini Ricerche, Florence, Italy

ABSTRACT

In single-arm clinical trials with survival outcomes, the Kaplan–Meier estimator and its confidence interval are widely used to assess survival probability and median survival time. The intent of this work is to write a code to compare the survival curve of a single treatment group to that of an historic control. Such is often the case in clinical phase-II trials with survival endpoints. Currently there is no specific PROC that can be applied to such analysis in SAS. There is availability of a library in R, to calculate the One Sample LogRank test (OneArmTTE). We are here presenting a calculation of the One sample LogRank test in SAS, which reproduce the exact results of the R library OneArmTTE. We also analyze through simulations the performance of such testing

INTRODUCTION

In single-arm clinical trials with survival outcomes, the Kaplan–Meier estimator and its confidence interval are widely used to assess survival probability and median survival time. The intent of this work is to compare the survival curve of a single treatment group to that of an historic control. Such is often the case in clinical phase-II trials with survival endpoints. While there are methods developed in R, such as library OneArmTTE, there is no procedure that can be applied in SAS.
In the present work, we develop a code in SAS which is built in order to reproduce what is done by the OneArmTTE library in R. We also propose simulations to better evaluate the performance of the methodology.

METHODOLOGY

Let us consider here a dataset where, for each patient, we have a time (time-to-event) and a censoring indicator.
We also have an historic control with a median time.
We assume a Weibull distribution for the survival curve.
The OneArmTTE in R calculates the Cumulative Hazard at a Landmark Timepoint
Given a dataset containing the time-to-event information, the function can get analysis results on the input trial data using several approaches for one-arm trial design with time-to-event endpoint, including one-sample log-rank test, Landmark Kaplan-Meier method and binary method which regards the survival of each subject at a landmark as a binary variable. The outputs of the function include the p-value of one sample log rank test, the historical survival rate at landmark and the survival rate estimate with confidence interval of landmark Kaplan Meier method.
Our SAS code reproduces the p-value given by the one sample log rank as calculated by the R function.

THEORETICAL METHOD

We calculate the observed events using d_i

$$O = \sum_{i=1}^{n} d_i$$

We then calculate the expected with ρ global rate of events in the Weibull distribution and Me median of reference

$$\begin{cases} \rho = \dfrac{\ln 2}{Me} \\ E = \displaystyle\sum_{i=1}^{n} t_i * \rho \end{cases} \Rightarrow E = \frac{\ln 2}{Me} \sum_{i=1}^{n} t_i$$

Starting from a Weibull with shape parameter k, historic median of reference Me, we calculate λ scale parameter as a function of Me and k

$$\lambda = \frac{Me}{\sqrt[k]{\ln 2}}$$

We then calculate the survival function $S(t)$ hypothesizing the Weibull distribution

$$S(t_i) = \exp\left\{-\left(\frac{t_i}{\lambda}\right)^k\right\} = \exp\left\{-\left(\frac{t_i}{\frac{Me}{\sqrt[k]{\ln 2}}}\right)^k\right\} = \exp\left\{-\left(\frac{t_i * \sqrt[k]{\ln 2}}{Me}\right)^k\right\}$$

We calculate the hazard function

$$H(t_i) = -\ln\big(S(t_i)\big) = -\ln\left(\exp\left\{-\left(\frac{t_i * \sqrt[k]{\ln 2}}{Me}\right)^k\right\}\right) = -\left\{-\left(\frac{t_i * \sqrt[k]{\ln 2}}{Me}\right)^k\right\} = \left(\frac{t_i * \sqrt[k]{\ln 2}}{Me}\right)^k = t_i^k * \frac{\ln 2}{Me^k}$$

The expected events are

$$E = \sum_{i=1}^{n} H(t_i) = \sum_{i=1}^{n} t_i^k * \frac{\ln 2}{Me^k} = \frac{\ln 2}{Me^k}\sum_{i=1}^{n} t_i^k$$

If $k = 1$

$$k = 1 \Rightarrow E = \frac{\ln 2}{Me^k}\sum_{i=1}^{n} t_i^k = \frac{\ln 2}{Me}\sum_{i=1}^{n} t_i$$

The test statistic is then

$$\boldsymbol{Z = \frac{O - E}{\sqrt{E}}}$$

CODE

The code reproducing the calculation is the following.
Firstly, we create the dataset with survival data. We assume a 12 months recruitment and a study duration of 3 years. We also assume that the only reason for censoring is end of study.

```
do subj=1 to 150;
c = rand("WEIBULL", shape, scale); /*Simulating survival time */
t = 36-rand("UNIFORM")*12; /*observation time*/
time=min(c,t);
if c<=t then censored=0; else censored=1 ;
output;
```

We create a status indicator and set the median of reference

```
data dataset1; set pop;
keep sim time status;
status=(censored=0);
time=time**&shape.;
run;

/*median of reference*/
%let median=(5+&j./10);
```

We then calculate the sum of the survival times

```
/*Sum up times and observed*/
ods exclude all;
ods output Summary=sums1;
proc means data=dataset1 sum; var time status; by sim; run;
```

The test statistic is calculated as

```
/*Test*/
data sums2;
set sums1;
observed=status_sum;
expected=time_sum*log(2)/&median.**&shape.;
Z=abs((observed-expected)/expected**.5);
   pvalue =2*(1-PROBNORM(abs(Z)));
keep sim observed expected Z pvalue sign;
if pvalue<0.05 then sign="Y"; else sign="N";
run;
```

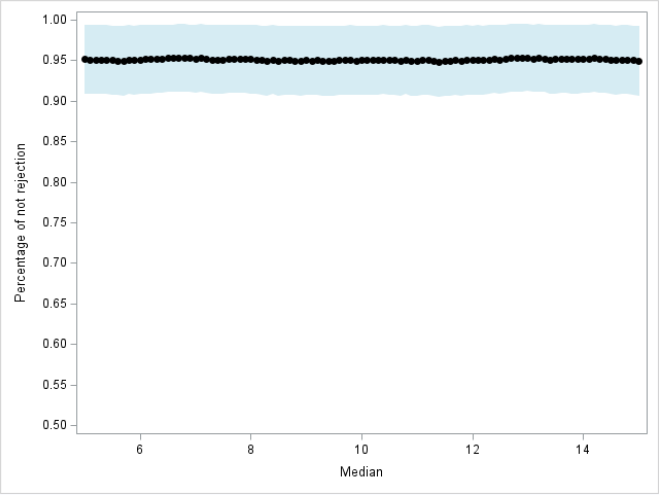
This code is reproducing exactly the same p-value generated by R with the library OneArmTTE.

For the respective R code see Appendix 1.

SIMULATIONS

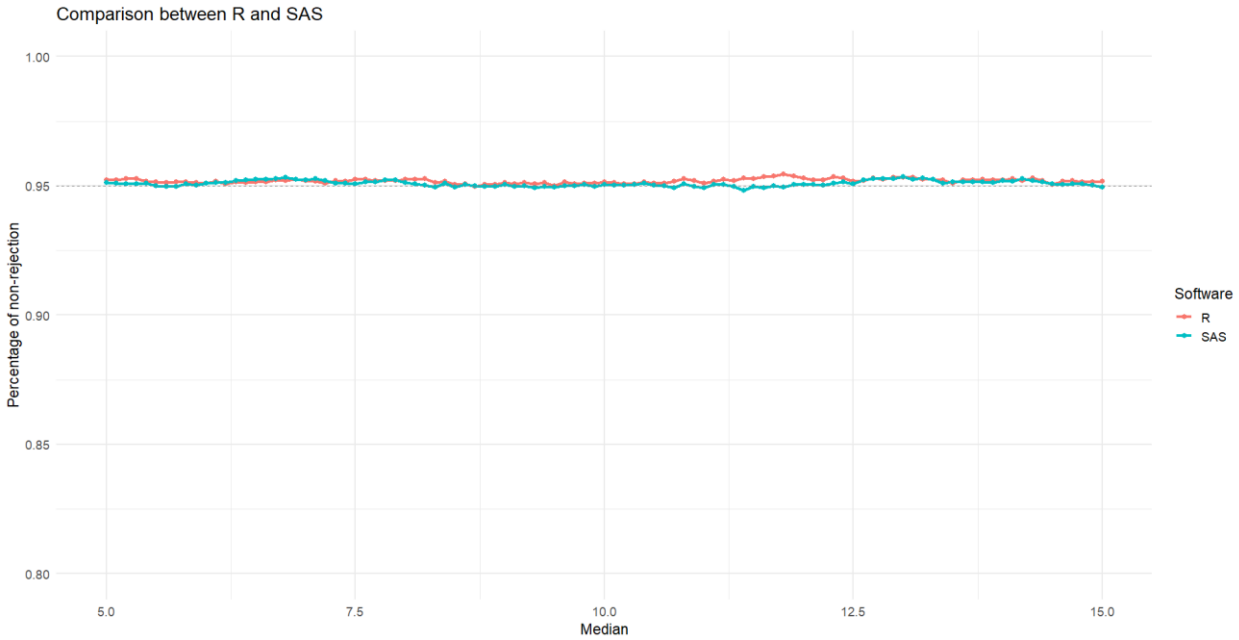
To check the performance of the SAS code, firstly we tried to simulate a case where the null hypothesis is true, varying the median of reference.
Given a certain median for the historical control, we simulated the survival dataset sampling from a Weibull with that median.
We then repeated the 10000 simulation for different value of the median ranging from 5 to 15.
As we observe in Figure 1 about the percentage of not rejection of null hypothesis with the 95% CI, the percentage of the p value >0.05 is stable around 95%, showing that the results are stable regardless of the median of reference.

Figure 1



Secondly, we repeated the previous analysis using the library OneArmTTE in R. We produced the following plot comparing R and SAS results

Figure 2

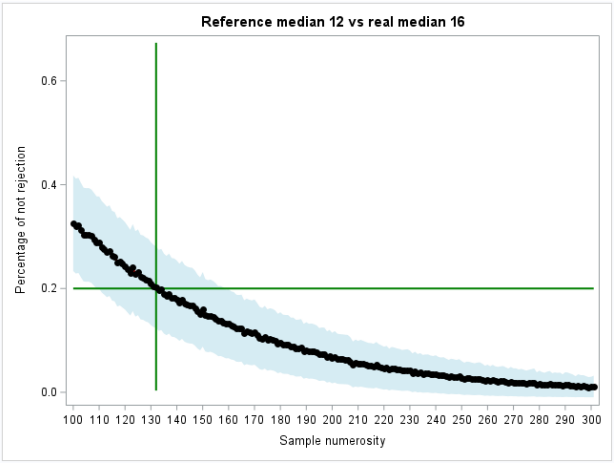


The plot shows that the SAS code reproduces the R code in the significance assessment of the One Sample Log Rank Test.

Regarding the stability of significance, it can be seen that the two software programs do not show significant differences and that in both cases no variability was obtained around the expected value of 0.95 depending on the median value.

Subsequently, we repeated the simulations varying the sample size.
We assumed a median of reference of 12 months and a true median of 16 months.
We calculated the sample size using PASS software. Assuming a power of 80% we obtain a sample size of 132 patients.
Using these two values for the medians, we simulated the statistics on 10000 datasets calculating the p-value, varying the sample size from 100 to 300.
We see the results in Figure 3.

Figure 3.



We see that around n=132, the percentage of non-significant p-values is around 20% as expected.

CONCLUSION

The present work is aimed to present a SAS code for the calculation of the One Sample Log Rank test
The code is reproducing the p-value from the log-rank test calculated in R by the library OneArmTTE.
We presented a set of simulations to show the performances of the test statistic.
In addition, a comparison between SAS code and PASS was done, obtaining the same results.
The present code can be used in any Phase 2 study with a time-to-event endpoint where there is only one sample and the reference median is taken from historical control.
The main uses for this test are:

- If done retrospectively, to evaluate the probability that a time-to-event under the null hypothesis is coming from a Weibull with the median of the historic control
- If done prior to the study, to calculate the sample size necessary to obtain a certain power level.

The main advantage in using the SAS code instead of the R library OneArmTTE is that the SAS code implements the possibility to extend the test with any shape parameter k while the R library only allows the case with $k = 1$.

REFERENCES

J. P. Klein, M. L. Moeschberger, "Survival Analysis", Second Edition, Springer, New York , 2005.
Andersen, P. K. , Borgan O., Gill, R. D. and Keiding, N. "Statistical Models Based on Counting Processes, New York, Springer-Verlag , 1993.
Breslow, N. E. , "Analysis of Survival Data under the Proportional Hazards Model", , International Statistics Review 43 (1975):45-48

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:
Author Name: Giulia Tonini
Company: Menarini Ricerche
Address: via Sette Santi 1, Firenze
Email: gtonini@menarini-ricerche.it

APPENDIX 1

R CODE

```
df<- the dataframe with simulated data used containing sim (simulation number
1:10000), time (time to event variable) and status (1=event, 0=no event)

results<-c()
 for (i in seq(5, 15, by = 0.1)) {
   for (j in 1:10000){
median.ctrl<-c(i,i,i)
eventRates.ctrl <- tibble::tibble(duration=c(36,0,0),rate=log(2)/median.ctrl)
step<-df[df$median==median.ctrl[1],]
step<-step[step$sim==j,]
step$censor<-prova$status
res <- capture.output(
  OneArmTTEAnalysis(step, eventRates.ctrl, landmark = (max(step$time)-0.01))
)
pval_logrank <- as.numeric(sub(".*: ", "", res[1]))
step2<-cbind(i,pval_logrank)
results<-rbind(results,step2)
 }}
```