SURVIVAL ANALYSIS FOR CLINICAL STUDIES

Katerina Langova

Department of Medical Biophysics, Faculty of Medicine and Dentistry, Palacky University Olomouc, Czech Republic
e-mail: langova@tunw.upol.cz

Aim: This paper focuses on the use of censored data in survival analysis. Survival analysis is used most frequently in the case of cancer patients when the study is finished and a number of individuals are still alive. The original article cited[2] was declared recently to be the most cited statistical study in the biomedical area. The goal of this paper is to explain the basic principles and methods involved. The way survival analysis processes data and interprets outputs is presented using the clinical data of oncological patients.

Methods and Results: Survival analysis is used to estimate survivor function from survival data, to compare survivor functions and to assess the relationship of explanatory variables to survival time. These methods were applied to the data of 176 patients with heamato-oncological diagnoses who had undergone bone marrow blood transplant.

Conclusion: It is very important to use appropriate methods when processing statistical data. Standard statistical procedures used for incomplete data could not provide the correct estimation.

INTRODUCTION

Of statistical methods applied to data evaluation in biomedical research, survival analysis has been proving to be increasingly significant since the 1950s. This method mines a maximum of relevant information from studies finished at a time when the collected data has not been completed. A typical example is a study in which the dependent variable is the survival or the length of life of cancer patients. However, from the literature and from biometrical experience, survival analysis has also been applied to materials, machinery and electrotechnical products as well (material fatigue) as human patients. This is attested to in the first two articles on this topic: the first author focused on cancer studies while the second on the lifespan of electronic tubes.

MATERIAL AND METHODS

The statistical department of the Institute of Medical Biophysics has been applying survival analysis for more than two decades. The impulse to begin using the method was a requirement for analysis of incomplete survival data obtained mostly from oncological patients. These data were collected by physicians (eg. in collaboration with a team of the 3rd Internal clinic headed by prof Ščudla) and haematooncologists (headed by prof Indrak). Generally survival analysis is a collection of methods that process the variable of "time to the occurrence of an event". This variable is also often called "time of survival" or briefly "survival". The "survival time" refers to a number of years, months, weeks or days from the beginning of the patient observance till the occurrence of an observed event (death as the rule). Most studies on survival analysis are terminated before the observed event (end-point) occurs for all subjects. This situation is called in survival analysis "censoring".

There can be three reasons for censoring:
1. the observed event does not occur for an object till the end of a study
2. an object was eliminated from study or decided to abandon the study
3. an object is lost from observation before the end of the study (eg. he died for the reason other than that of the event under study)

Biomedical examples which can be analyzed by the survival method:
- length of the patient remission with acute leukaemia till relapse occurrence or the time to death from the beginning of the remission.
- survival of a female patient with mammary carcinoma with various immunohistochemical response.
- time of effect of analgesic after kidney surgery in infants and toddlers to the application of opiates in the case of pain manifestation
- time to re-infection in case of patients suffering from sexually transmitted disease.

After finishing a clinical study we obtain input data for subsequent statistical processing. Usually there is a group of patients with an observed diagnosis available.

The group consists of three subgroups of individuals:
1. patients for which an event occurred (e.g. death) – final data
2. patients for which the event did not occur – censored data
3. patients who were eliminated from a study for some reason or who cannot be observed further – also censored data.

There are two items of data for each patient:
1. survival time (means period of individual observation; also used for censored data)
2. reason why the individual observation was terminated – coded as 1 for the observed event, 0 for patients from groups 2 and 3.

In our study the subject of analysis is the data from 176 patients with heamatooncological diagnoses who have undergone bone marrow blood transplant and then a number of patients suffering from acute graft-versus-host disease (GvHD) which is a common side-effect of the transplant.

Data of the study are in Table 1:

Table 1. Input data for survival analysis

Group	OS	status
0	4.3	1
0	10.2	1
1	3.6	1
0	27.6	1
.	.	.
.	.	.
.	.	.
1	21.6	1
1	29.1	0
0	29.3	0
1	18.0	1

group 1 comprised patients with acute GvHD; group 0 consists of patients without acute GvHD; OS is overall survival in months; status 1 – patient died, 0 patient lives.

The goals of survival analysis are to:
1. estimate and interpret survivor function from survival data
2. compare survivor functions
3. assess the relationship of explanatory variables to survival time.
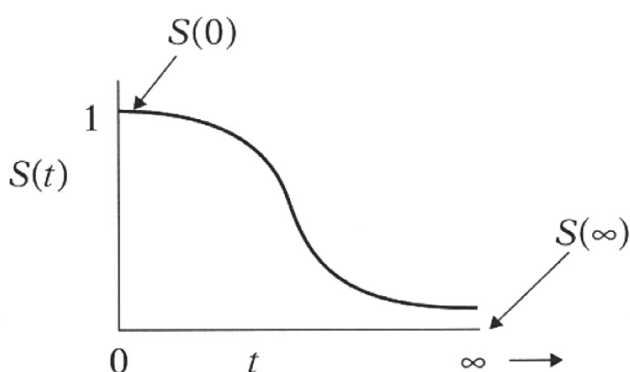
Graph 1. Theoretical shape of survivor function

1. Survivor function

Survivor function describes the probability that a person lives longer than a specified time t since the beginning of observation ($S(t) = P(T > t)$, T is a nonnegative random variable which spans the time from the beginning of an observation to the occurrence of the observed event - survival time).

2. Comparison of survivor functions

A common problem in clinical studies is to compare two or more survivor functions. There are a few statistical tests for such a comparison. The most used tests are the log rank test and generalized Wilcoxon test (also called Breslow test).

The log rank test is in fact a chi-squared test for a large sample. The log rank statistic compares the observed with an expected number of events. The expected number of events is calculated by the method assuming that the null hypothesis is true. The null hypothesis assumes that the compared curves are the same. The comparison is performed at every time point the observed event occurred.

The Wilcoxon test does not work with real times to an occurrence of the observed event but only with a rank of events'occurrence in two groups. Each ranked item of one group is compared with all ranked values from a second group. This eliminates the effects of extreme times from input groups.

3. The Cox proportional hazards model

The Cox proportional hazards (PH) model is used to test the effect of a number of variables or factors on survival and hazard an observed event occurrence. This multiple regression method reveals statistically significant factors and also quantifies the effect of all factors on survival.

A sufficient number of observations, no correlation between two or more independent variables and independence of observations are general assumptions of regression methods. In addition the Cox PH model further requires that the effects of the predictor variables are constant over time.

Input data for the Cox PH model:
1. time to occurrence of an observed event (dependent variable)
2. reason for terminating observation
3. independent variable predictors – factors (can be both quantitative and categorical)

The Cox PH model (for p independent variables X_1, ..., X_p is described by the equation:

$$h(t)/h_0(t) = e^{(B_1X_1+B_2X_2+...+B_pX_p)},$$

where $h(t)$ is the hazard function, $h_0(t)$ is called the baseline hazard function (the expected hazard without any effect of the considered factors), e is a base of the natural logarithm, B_1, ..., B_p regression coefficients.

The expression $h(t)/h_0(t)$ is called the hazard ratio (HR) and indicates a growth or decrease of hazard caused by an effect of factors X_1, ..., X_p.
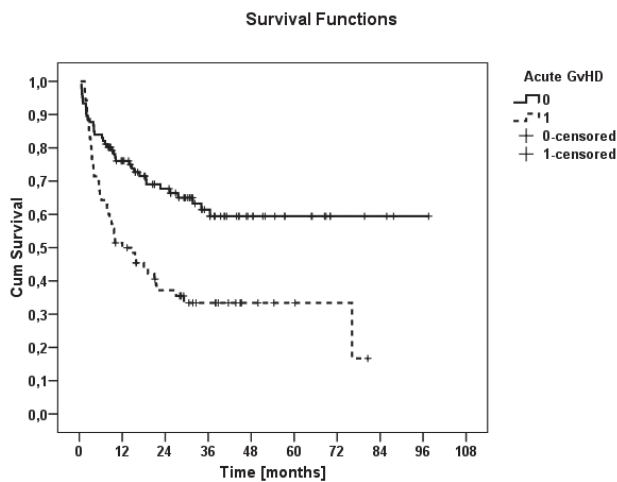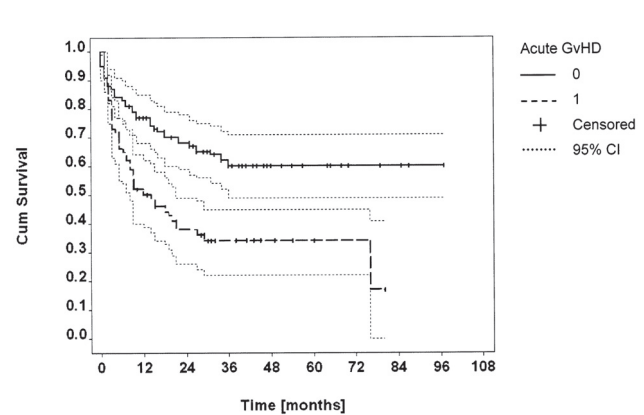
Graph 2. The real shape of survivor functions for the two groups.

Graph 3. The real shape of survivor functions for the two groups with 95% confidence intervals.

Calculation of the Cox regression equation is included in most statistical software applications. The coefficients B_1, ..., B_p and their standard errors SE are estimated using complex mathematical methods and the significance of these coefficients in the Cox PH model is tested with the Wald statistic. If the Wald test statistic probability is 0.05 or lower we consider the independent variable as an effective predictor of the observed event. The expression e^{B_i} (Exp(B_i)), $i = 1$, ..., p estimates a percentage change of the observed event hazard (e.g. death) for unit change of the independent variable. In Cox PH model output we detect in particular if the examined predictors are statistically significant and if they are then we interpret HR for them.

RESULTS AND DISCUSSION

1. Survivor function

In practice, survivor function is estimated using the Kaplan-Meier method (K-M method). $S(t)$ denotes a theoretical survivor function that holds true for the whole population and $\hat{S}(t)$ means a survival function for a given sample or study group. Values of the function $\hat{S}(t)$ are calculated in all times when the observed event occurred for any patient.

The Kaplan-Meier method provides very good estimations of survival probabilities also in such cases when the survival distribution is unknown. This method assumes that the pattern of censoring is independent of the survival times.

The real shape of survivor function for both groups observed in our study is shown in graph 2. The horizontal axis indicates time. The vertical axis displays survival probability of a time ranging from 0 to 1.

The SPSS application is not able to display the 95% confidence interval for survivor functions. To solve this problem we cooperated with a programmer and designed and implemented a new software program in our department. The program is written in C++ language using MFC libraries and was developed in Visual Studio environment from Microsoft. An output from this program is shown in graph 3.

When survival probability is calculated, the corresponding statistical method will also provide us with medians and means for survival time. Median for survival time is the time at which the survival probability equals 0.5. Means for survival time is defined as the area under the survivor function curve.

Table 2. Means and medians for survival time; medians for the group without acute GvHD could not be calculated because the survival probability at the end of observation was more than 0.5.

Group	Mean				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
0	63.067	4.522	54.203	71.931	.	.	.	.
1	32.121	4.133	24.020	40.222	12.033	4.278	3.648	20.417
Overrall	51.270	3.754	43.912	58.628	34.098	15.132	4.440	63.756

Table 3. Overall comparisons. Almost identical results of both statistical tests (in both cases p < 0.05) show us that the patients with acute GvHD have a worse survival.

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	14.508	1	.0001
Breslow (Generalized Wilcoxon)	11.984	1	.0005

Table 4. Factors selected for the regression model

	Factors (variables)	Categories or units
Categorial variables	Male host	Yes – no
	Male donor	Yes – no
	Acute leucaemia diagnosis	Yes – no
	Related donor	Yes – no
	ATG or Campath 1 H (drugs) in a preparatory mode	Yes – no
	Graft rejection after transplant	Yes – no
	Presence of IL6174 G allele in donor	Yes – no
	Presence of IL6174 G allele in host	Yes – no
	Presence of IL6nt565 G allele in donor	Yes – no
	Presence of IL6nt565 G allele in host	Yes – no
	Presence of MCP 12076 T allele in donor	Yes – no
	Presence of MCP 12076 T allele in host	Yes – no
	Presence of MCP 12518 G allele in donor	Yes – no
	Presence of MCP 12518 G allele in host	Yes – no
Quantitative variables	Age of host	years
	Age of donor	years
	CD34+ (number of stem cells)	10^6 per 1 kg

Table 2 shows the output from the SPSS program with calculated estimates, standard errors and 95% confidence intervals of means and medians (if possible) for both observed groups and the overall group.

From the table it is obvious that the patients with acute GvHD had lower values of survival time.

2. Comparison of survivor functions

The output from the SPSS program in table 3 shows level of significance from a corresponding comparison of the survivor function curves of both observed groups.

3. Cox PH model

In our study we evaluated the effect of selected factors on survival and hazard of an observed event occurrence (death). Table 4 displays an overview of categorial and quantitative factors using the Cox PH model:

There were 6 models designed with the stepwise regression model method (this is a method of a gradual addition of factors to a model) – see steps 1 to 6. *B* coefficients and HR (a change of hazard for a unit change of a factor value) were estimated for the factors assigned to the model.

Table 5 presents the following statistics for each variable:

B is a coefficient estimation in the Cox PH model

SE is a standard error of *B*

Wald is a Wald statistic = $(B/SE)^2$; it has a chi-square distribution and tests the hypothesis $B = 0$.

df are degrees of freedom

Sig. is a significance for the Wald statistic

$Exp(B)$ – e^B estimates a hazard change for a unit variable step

Cox PH model summary:

If a patient takes ATG or Campathe 1 H (drugs) in a preparatory mode then the death hazard after the bone marrow transplant decreases 0,373 times. (95 % CI 0.22–0.633)

If the variable CD34 + (number of stem cells in 1 kg) increases by 10^6 then the death hazard after the bone marrow transplant decreases 0,94 times. (95 % CI 0.887–0.998)

If a graft rejection occurred after the transplant then the death hazard after the bone marrow transplant increases 3,25 times. (95% CI 1.11–9.54)

Table 5. Variables in the equation

		B	SE	Wald	df	Sig.	Exp(B)	95.0 % CI for Exp(B)	
								Lower	Upper
Step 1	chronickáGvHD	−1.166	.286	16.602	1	.000	.311	.178	.546
Step 2	akutníGvHD	1.016	.225	20.316	1	.000	2.763	1.776	4.298
	chronickáGvHD	−1.346	.290	21.537	1	.000	.260	.148	.460
Step 3	ATGneboCampath1 Hvpřípravnémrežimu	−.999	.267	13.942	1	.000	.368	.218	.622
	akutníGvHD	.867	.228	14.516	1	.000	2.380	1.524	3.718
	chronickáGvHD	−1.428	.289	24.343	1	.000	.240	.136	.423
Step 4	ATGneboCampath1 Hvpřípravnémrežimu	−.968	.269	13.000	1	.000	.380	.224	.643
	akutníGvHD	.952	.232	16.857	1	.000	2.592	1.645	4.084
	chronickáGvHD	−1.521	.293	26.993	1	.000	.218	.123	.388
	přítomnostIL6174 Galelyudárce	.732	.329	4.960	1	.026	2.079	1.092	3.961
Step 5	ATGneboCampath1 Hvpřípravnémrežimu	−.931	.269	11.981	1	.001	.394	.233	.668
	rejekceštěpupoTKB	1.170	.546	4.588	1	.032	3.222	1.104	9.399
	akutníGvHD	1.047	.241	18.931	1	.000	2.848	1.777	4.563
	chronickáGvHD	−1.504	.295	26.061	1	.000	.222	.125	.396
	přítomnostIL6174 Galelyudárce	.819	.333	6.033	1	.014	2.268	1.180	4.360
Step 6	ATGneboCampath1 Hvpřípravnémrežimu	−.986	.270	13.328	1	.0003	.373	.220	.633
	CD34x106kg	−.061	.030	4.145	1	.042	.940	.887	.998
	rejekceštěpupoTKB	1.180	.549	4.622	1	.032	3.253	1.110	9.536
	akutníGvHD	1.190	.247	23.284	1	< 0.0001	3.287	2.027	5.331
	chronickáGvHD	−1.632	.300	29.504	1	< 0.0001	.196	.109	.352
	přítomnostIL6174 Galelyudárce	.889	.337	6.959	1	.008	2.433	1.257	4.709

If an acute GvHD developed after the transplant then the death hazard after the bone marrow transplant increases 3,29 times. (95 % CI 2.03–5.33)

If chronic GvHD developed after the transplant then the death hazard after the bone marrow transplant decreases 0,2 times. (95 % CI 0.11–0.35)

The presence of IL9174 G allele in a donor increases the death hazard after the bone marrow transplant 2,433 times. (95 % CI 1.26–4.71)

CONCLUSIONS

We wish to issue a caveat on the dangers of statistical misinterpretation of data when applying standard statistical procedures in the case of incomplete (censored) data. For example a computation of means or medians without respecting data incompleteness would lead to an underestimation of the reality (in terms of real patient survival).

REFERENCES

1. Cox DR, Oakes D. Analysis of Survival Data. Chapman & Hall, USA 1998
2. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Amer Statist Ass 1958; 53:457-481.
3. Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated Data. Springer, New York 1997
4. Kleinbaum DG. Survival Analysis. Springer-Verlag, New York 1996
5. Snedecor GW, Cochrane WG. Statistical Methods. Ames, The Iowa State University Press, 1967