Survival Analysis

Harald O. Stolberg[1,2]
Geoffrey Norman[3]
Isabelle Trop[4]

Received November 17, 2004; accepted November 23, 2004.

Series editors: Nancy Obuchowski, C. Craig Blackmore, Steven Karlik, and Caroline Reinhold.

This is the 19th in the series designed by the American College of Radiology (ACR), the Canadian Association of Radiologists, and the *American Journal of Roentgenology*. The series, which will ultimately comprise 22 articles, is designed to progressively educate radiologists in the methodologies of rigorous clinical research, from the most basic principles to a level of considerable sophistication. The articles are intended to complement interactive software that permits the user to work with what he or she has learned, which is available on the ACR Web site (www.acr.org).

Project coordinator: G. Scott Gazelle, Chair, ACR Commission on Research and Technology Assessment.

Staff coordinator: Jonathan H. Sunshine, Senior Director for Research, ACR.

[1]Department of Radiology, McMaster University Medical Centre, 1200 Main St. W, Hamilton, ON, L8N 3Z5 Canada.

[2]Deceased.

[3]Department of Educational Research, Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, ON, Canada.

[4]Department of Radiology, Hospital St.-Luc, 1058 St-Denis St., Montreal, QC, H2X 3J4 Canada. Address correspondence to I. Trop.

AJR 2005;185:19–22

0361–803X/05/1851–19

The breadth of radiology research is expanding. Previously, a large proportion of radiology research projects were observational studies. Increasingly, research now involves groups of patients to whom specific interventions are administered in a randomized fashion. Analysis of data obtained from these experimental studies varies, depending on the end point of interest. Research protocols that are designed to evaluate the interval between entry of a patient into the study and the time until the event of interest are referred to as time-to-event studies, a form of follow-up study [1]. The event may be death in a diagnostic study of cancer or a progression of various chronic disease entities to a defined stage. In interventional studies, such as vascular and neuroradiologic procedures, the fate of grafts, stents, and other devices may be followed through time. Survival analysis, also called "life table" analysis, refers to the methodology of analysis of data gathered in such protocols. Survival analysis, then, is the topic of this article [2].

Overview

Under ideal circumstances, a study would enroll all its subjects simultaneously and follow them either for a fixed period of time or until they all reach some end point, such as recovery or death. However, more commonly, studies require a large number of subjects or look at relatively rare conditions, and so must enter subjects over a period of several months or even years. When the study finally ends, the subjects will have been followed for varying lengths of time, during which a number of different outcomes have to be considered: the event has not yet occurred (outcome A), some patients are lost to follow-up (outcome L), or the event has occurred (an example of the event or end point is death) (outcome D).

Figure 1 shows how we can illustrate these different outcomes, indicating what happened to the first 10 patients in a study. Subjects A, C, D, and F died during the trial; they are labeled "D" for dead. Subjects B, G, and I were lost to follow-up, hence the label "L," at various times after they started the drug. The other subjects, E, H, and J (labeled "C"), were still alive at the time the trial ended. These last three data points are called "right-censored." Subjects are considered "censored" when their data are incomplete. They are said to be right-censored because they have been followed to the end of the study (the "right-hand part" of the graph), but the outcome of interest has not occurred to them. To be more quantitative about the data, Table 1 shows how long each person was in the study and what the outcome was.

The Kaplan-Meier Approach to Survival Analysis

To do a survival analysis, we must figure out how many people survive for at least 1 year, for at least 2 years, and so on, in what

TABLE 1: Outcomes of the First 10 Subjects

Subject	Length of Time in Trial (months)	Outcome
A	61	Died
B	111	Lost
C	29	Died
D	46	Died
E	92	Censored
F	22	Died
G	37	Lost
H	76	Censored
I	14	Lost
J	45	Censored

Note—Reprinted with permission from [4].

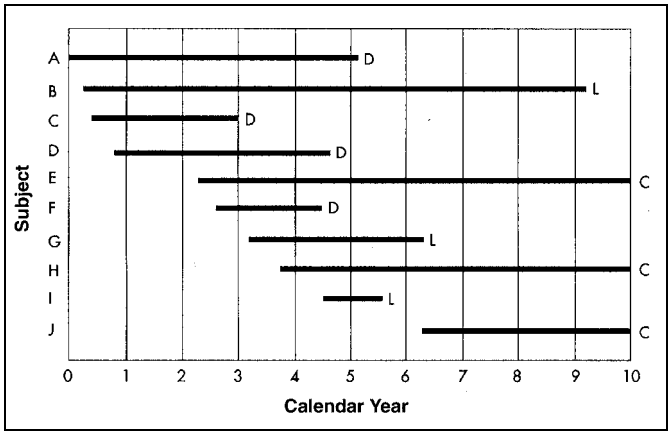

Fig. 1—Entry and withdrawal of subjects in a 10-year study. (Reprinted with permission from [4])

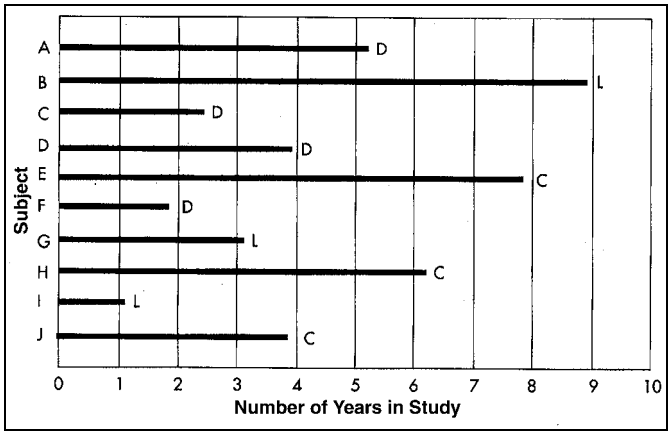

Fig. 2—Figure 1 redrawn so all subjects have a common starting date. (Reprinted with permission from [4])

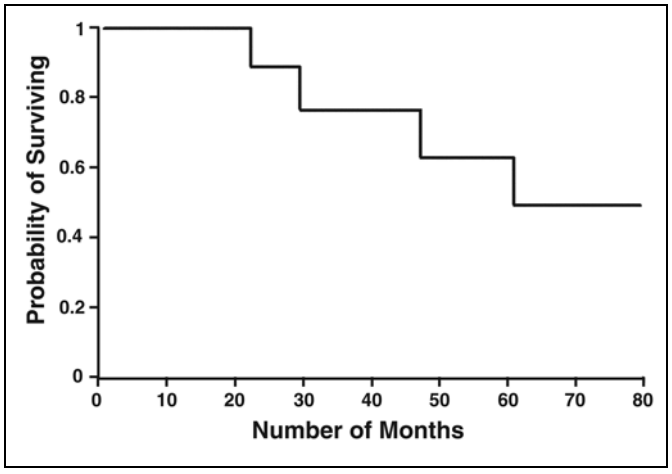

Fig. 3—Survival curve for data in Table 2.

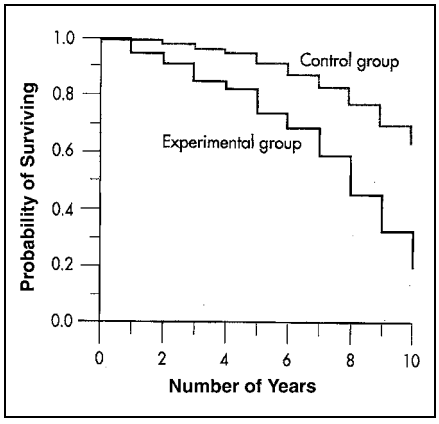

Fig. 4—Survival curves for both groups in study of patients with intramural hematoma of the aorta. (Reprinted with permission from [4])

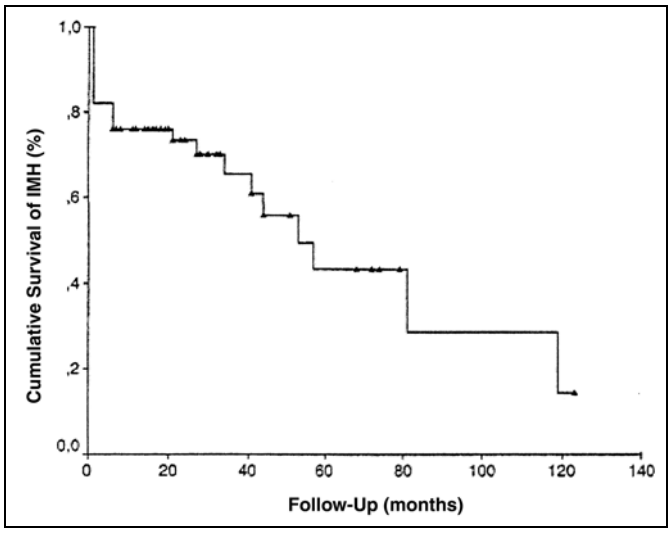

Fig. 5—Probability of survival after aortic intramural hematoma (IMH) in 66 study patients. Small triangles indicate censored cases.

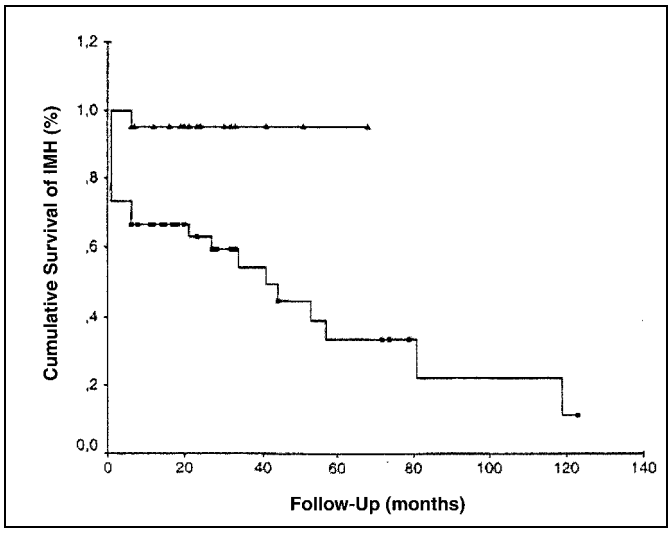

Fig. 6—Cumulative survival of patients with intramural hematoma (IMH) with (experimental group) and without (control group) treatment with β-blockers. Upper curve (*triangles*) indicates treated patients; small squares indicate censored cases. Difference between two subgroups was statistically significant ($p = 0.004$).

is called a "life table" technique. There are two ways to go about calculating a life table: the actuarial approach and the Kaplan-Meier approach [3]. The Kaplan-Meier approach is far more common in medical literature, so we will describe it.

The first step involves redrawing the graph, so that all the people appear to start at the same time. Figure 2 shows the same data as Figure 1; however, instead of the *x*-axis being Calendar Year, it is now Number of Years in Study. The lines are all the same length as in Figure 1; they have just been shifted to the left so that they all begin at time 0.

The Kaplan-Meier approach uses the exact time of death in the calculation of survival. It also computes the survival function only when an outcome occurs. To show how this is done, let us use the data for the 10 subjects in Table 1. The first step is to rank-order the length of time in the trial and flag which entries reflect the outcome of interest (death in this case) and which are due to withdrawal or censoring. We have done this by putting an asterisk after the data for subjects who were lost to follow-up or were censored by the termination of the study:

14* 22 29 37* 45* 46 61 76* 92* 111*

This data set would generate a life table (Table 2) with only four rows, one for each of the four patients who died.

One person was lost to follow-up before the first person died, so the number of remaining patients at risk at 22 months is only nine. Death rate, survival rate, or any other statistical estimate is calculated on the basis of the population at risk (Table 2). At 46 months, two people had died and three were lost to follow-up, so the number of patients at risk is five, and so on. This little data set would generate a survival curve like that shown in Figure 3 except for fewer steps.

Comparing Two (or More) Groups with the Log-Rank Test

Although the survival curve shown in Figure 3 tells us what happened to patients over time, we often want to compare two or more groups of patients—for example, patients with different kinds of stents, or patients who were screened (experimental group) versus patients who were not screened (control group). So we will create an expanded survival table with 250 experimental subjects and 250 control patients. These data are presented in Figure 4. This graph shows that the survival curve for the treatment group dropped at a faster rate than that for the control group. But is the difference statistically significant?

The best approach for evaluating whether the difference is indeed significant is to use the Mantel-Cox log-rank test, which is a modification of the Mantel-Haenszel chi-square test [4]. This test is a powerful method for analyzing data when the time to the outcome is important; it deals with censored data and differential length of follow-up of different subjects. As with most chi-square tests, the log-rank test compares the observed number of events with the number expected, under the assumption that the null hypothesis of no group differences is true. That is, if there were no differences between the groups, then at any interval, the total number of events should be divided between the groups roughly in proportion to the number of subjects at risk. The test determines how much the observed event rate differs from the expected rate.

The Cox Proportional Hazards Model

A more sophisticated method of analysis commonly used, which examines the difference in the survival curves while also accounting for other variables (covariates), is the Cox proportional hazards model [5]. Unlike the log-rank test, the proportional hazards model allows adjustment for any number of covariates, whether they are discrete (e.g., the technique used [CT or MRI]) or continuous (e.g., age or serum electrolyte level), and then computes a test for each, including, of course, a statistical test of the difference overall between the treatment and control groups. Both survival and hazard functions can refer to outcomes other than death. In the Cox model, this hazard is assumed to be separable into a product of one function that depends on time and another function that captures all the other variables including, specifically, the relative difference between treatment and control groups.

No matter which form of survival analysis statistical test is used, four assumptions must be met:

- Each person must have an identifiable starting point. All subjects should enter the trial at the same time in the course of their illness. Using diagnosis as an entry time can be problematic, because people may have had the disorder for varying lengths of time.
- A clearly defined and uniform end point is required. This is not a problem if the end point is death, but it can be a problem if the end point is recurrence of disease.
- The reasons that people drop out of the study cannot be related to the outcome. If persons have dropped out because they can no longer travel to their scheduled appointments as a result of the worsening of symptoms of the disease under study, the chances of survival could be seriously overestimated. Otherwise, any changes we see may be due to these secular changes, rather than the intervention.
- Diagnostic and treatment practices must not change over the life of the study.

We have said that survival or life table analysis allows us to look at how long people are in one state (e.g., life) followed by a discrete outcome (e.g., death). This analysis can handle situations in which the people enter the trial at different times and are followed up for varying periods; it also allows us to compare two or more groups [4]. The methods of life table (survival) analysis are increasingly used in diagnostic imaging research in recent years, and we therefore offer a recent review of a relevant research study [6].

This multicenter study evaluated patients with intramural hematoma of the aorta and hospital admission less than 48 hr after onset of initial symptoms. Patients were enrolled between January 1994 and December 2000 after confirmation of intramural hematoma on two imaging studies (transesophageal

TABLE 2: Kaplan-Meier Life Table Analysis of the Data in Table 1

Time (months)	No. at Risk	No. of Deaths	Death Rate	Survival Rate	Cumulative Survival Rate
t	R_t	D_t	q_t	p_t	P_t
22	9	1	0.1111	0.8889	0.8889
29	8	1	0.1250	0.8750	0.7778
46	5	1	0.2000	0.8000	0.6222
61	4	1	0.2500	0.7500	0.4667

Note—Reprinted with permission from [4].

echocardiography, CT, or MRI). Sixty-six patients were consecutively enrolled over the course of 7 years. They were subjected to medical treatment in an ICU setting and surgical treatment if indicated (criteria for surgical intervention are available in the original article). Follow-up of these patients ranged from 6 to 123 months and included outpatient visits and CT 6 months after the event and yearly thereafter.

From the raw data collected from 66 patients, a Kaplan-Meier curve was built (Fig. 5). Dissecting Figure 5, we obtain the following information: survival is set at 100% at the beginning of the study, when patients initially present to the emergency department. Each ladder step indicates a drop in survival—that is, the death of a patient because that was the event defined as the main outcome. A rapid decline ensues because close to 20% of patients die in the acute phase. The first loss of information occurs around 6 months, when the first follow-up is scheduled. The triangles indicate censored data, and the figure shows that at 20 months, 12 patients have already been censored. Figure 5 shows that the drop in survival is faster in the initial months after intramural hematoma: the curve drops faster between 20 and 60 months than later in the study.

Differential survival of subgroups of the study was assessed using the log-rank test. The resulting Kaplan-Meier curves obtained from comparison of patients who received oral β-adrenergic receptor blockers (experimental group) and those who did not (control group) are displayed in Figure 6. Visual analysis easily reveals that patients taking β-blockers (upper curve) enjoyed much greater survival than patients who did not receive the medication (lower curve). In fact, the upper curve shows that only one patient died early in the study, and that subsequently all patients from whom information is available are still alive. However, many censored data points are seen, but there is no reason to believe these patients have died without knowledge of the study's investigators, which would falsely lead to the conclusion that β-blockers have a protective effect. The log-rank test performed on these two subgroups of patients revealed important information that was embedded in the initial Kaplan-Meier curve (Fig. 5) and could not have been obtained had it not been for this separate analysis.

Conclusion

In this article, we address life table and survival analysis and describe life table techniques such as the Kaplan-Meier approach. For the comparison of two or more groups, we describe the Mantel-Cox log-rank test. Finally, we discuss the Cox proportional hazards model, which examines the difference in the survival curves and also accounts for other variables (covariates). These statistical methods allow one to work with nontraditional units of analysis: person–time rather than person only. These tools are seen increasingly in the research literature and are gaining popularity in radiology research. These methods of data analysis have potential applications in many fields of radiology, most notably in the analysis of screening techniques and interventional studies.

Acknowledgments

Sadly, Dr. Stolberg passed away last January. His determination and enthusiasm were key in seeing this project to completion, and we are indebted to him for all he accomplished.

We express our appreciation and gratitude to Monika Ferrier for her patience and support, which is always a difficult task with several authors. She has kept us on track and prepared the manuscript.

References

1. Norman GR, Streiner DL. *PDQ statistics*, 2nd ed. St. Louis, MO: Mosby, 1997
2. Altman DG, Machin D, Bagant TN, Gardner MJ. *Statistics with confidence*, 2nd ed. London, UK: BMJ Books, 2000
3. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Statist Assoc* 1958; 53:457–481
4. Norman GR, Streiner DL. *Biostatistics: the bare essentials*, 2nd ed. Hamilton, ON: B. C. Decker, 2000
5. Cox DR. Regression models and life tables. *J Roy Statist Soc* 1972; 34:187–220
6. Von Kodolitsch Y, Csosz SK, Koschyk DH, et al. Intramural hematoma of the aorta: predictors of progression to dissection and rupture. *Circulation* 2003; 107:1158–1163

The reader's attention is directed to earlier articles in the Fundamentals of Clinical Research series:

1. Introduction, which appeared in February 2001
2. The Research Framework, April 2001
3. Protocol, June 2001
4. Data Collection, October 2001
5. Population and Sample, November 2001
6. Statistically Engineering the Study for Success, July 2002
7. Screening for Preclinical Disease: Test and Disease Characteristics, October 2002
8. Exploring and Summarizing Radiologic Data, January 2003
9. Visualizing Radiologic Data, March 2003
10. Introduction to Probability Theory and Sampling Distributions, April 2003
11. Observational Studies in Radiology, November 2004
12. Randomized Controlled Trials, December 2004
13. Clinical Evaluation of Diagnostic Tests, January 2005
14. ROC Analysis, February 2005
15. Statistical Inference for Continuous Variables, April 2005
16. Statistical Inference for Proportions, April 2005
17. Reader Agreement Studies, May 2005
18. Correlation and Regression, July 2005