integers, existentially, and their parts only live as long as *they* live. Still, *in* them, we can call parts, parts.—But when, without circumlocution or disguise, I thus come over to your views, I insist that those of you who applaud me (if any such there be) should recognize the obligations which the new agreement imposes on yourselves. Not till you have dropped the old phrases, so absurd or so empty, of ideas "self-compounding" or "united by a spiritual principle"; not till you have in your turn succeeded in some such long inquiry into conditions as the one I have just failed in; not till you have laid bare more of the nature of that altogether unique kind of complexity in unity which mental states involve; not till then, I say, will psychology reach any real benefit from the conciliatory spirit of which I have done what I can to set an example.

Has the day passed in American psychology when its parts are just parts, its divisions just divisions, and the search for integers has been abandoned? Is the apartness so characteristic of contemporary psychology the consequence of the apartness so many people feel in their own lives? Is there no center that can hold psychology together? Must we resign ourselves to a modern psychology that is completely a reflection of modern living? Was our fall from earthly grace signaled when psychology labeled William James and John Dewey as philosophers, *only* philosophers?

I could have served the purposes of this occasion better if I had, as I initially fantasized, read you excerpts from the writings of these two giants. I ask your pardon for uncourageously succumbing to the demand characteristic of this occasion.

REFERENCES

Hiigard, E. (1978). Prefatory remarks to William James' presidential address to the American Psychological Association. In E. Hilgard (Ed.), *American psychology in historical perspective: Addresses of the presidents of the American Psychological Association.* Washington, DC: American Psychological Association.

James, W. (1978). Knowing things together. In E. Hilgard (Ed.), *American Psychology in Historical Perspective: Addresses of the Presidents of the American Psychological Association.* Washington, DC: American Psychological Association.

Plato. (1948). *Euthyphro, apology, and crito. Phaedo. The death scene.* New York: Bobbs-Merrill (1948).

Skinner, B. F. (1948). *Walden two.* New York: MacMillan.

American Journal of Community Psychology, Vol. 21, No. 2, 1993

Charting the Process of Change: A Primer on Survival Analysis[1]

Douglas A. Luke[2]
Michigan State University

Survival analysis is a powerful and useful technique for understanding qualitative change. This article provides a practical, nontechnical introduction to the use of survival analysis for social scientists. Important issues in using survival analysis are discussed, including research design, data preparation and management, and data analysis. Attendance data from a self-help organization are used to illustrate common survival analysis tasks such as describing the overall survival and hazard functions, examining covariate effects, and modeling the form of the hazard function over time. An appendix that discusses the strengths and weaknesses of current survival analysis computer programs is included.

KEY WORDS: survival analysis; self-help.

> It was ordained at the beginning of the world that certain signs should prefigure certain events.
>
> —Cicero, *De Divinatione*

Editor's note: Edward Seidman served as action editor for this article while serving as Associate Editor for Methodology.

[1]I thank William Davidson, Susan Englund, Bruce Rapkin, Kurt Ribisl, and three anonymous reviewers for their helpful comments on earlier drafts. The example data presented here were collected with the support of an NIMH grant (MH37390) awarded to Julian Rappaport and Ed Seidman.

[2]All correspondence should be addressed to Douglas Luke, Michigan State University, Department of Psychology, 129 Psychology Research Building, East Lansing, Michigan 48824.

The measurement of change is a fundamental task for social researchers. The etiology of a psychological disorder, the length of time it takes for a preventive intervention to take effect, how long those effects last, period of time until a battered spouse seeks legal intervention, the course a self-help group takes from birth to death, prediction of when a person may become violent—these are all examples of common social and psychological research questions that cannot be understood without measuring and understanding change.

Although many psychological theories define change as a series of incremental shifts in some variable that can be measured quantitatively, much, if not most change is actually qualitative in nature (Widaman, 1991). Change is seen as an event that marks the shift from one state to another; it is transformational rather than incremental. This viewpoint of change as transformation can be seen explicitly in applied psychological work on change in organizations (Quinn & Cameron, 1988), the life-cycles of behavior settings (Wicker & King, 1988), transformational processes in person-environment relations (Stokols, 1988), and in the more applied side of developmental psychology (Sroufe & Jacobvitz, 1989).

The purpose of this paper is to provide a practical introduction to survival analysis, one of the most useful statistical techniques for analyzing qualitative change. Despite its fundamental importance, change remains a neglected concept in psychology (Katz, 1980). This may be due in part to the conceptual, empirical, and statistical problems inherent in the measurement of change (Cronbach & Furby, 1970; Willett, 1987). However, it may also be due to researchers' reliance on a small set of established statistical tools, and an unfamiliarity with newer techniques designed to assess change (Collins, 1991). We are all familiar with the "Rule of the Hammer," and we know it is a mistake to define our basic theories based on the limitations of our instruments. Thus, this paper is an attempt to provide information about a relatively new tool that allows a more flexible approach to understanding change.

As the title suggests, this paper is designed to help get the interested reader up to speed as quickly as possible on how to use survival analytic techniques. There are already a number of good statistical treatments of survival analysis for those who want a more in-depth or technical understanding (Allison, 1984; Blossfeld, Hamerle, & Mayer, 1989; Kalbfleisch & Prentice, 1980; Lawless, 1982; Yamaguchi, 1991). However, most of these articles assume a high level of statistical expertise, and few if any provide enough practical information about how to design survival analysis studies, collect the data, and use available software to perform

the analyses.[3] For this reason, I spend most of my time in this paper introducing the techniques of survival analysis in a nontechnical manner, using real data to provide numerous grounded examples. Many advanced issues will be treated cursorily; wherever possible I refer the reader to more in-depth literature discussing these topics.

I begin by briefly defining important survival analysis terms, providing some historical background, and showing how survival analysis is relevant to many common research situations. Also, I describe a study that provides the data for the survival analysis examples appearing throughout the paper. The main section of the paper provides a detailed description with examples of how survival analysis might be used by the typical researcher. Advanced and more technical topics are then briefly covered. Finally, the paper includes an appendix that discusses computer programs that are available for survival analysis.

DEFINITION AND BACKGROUND

Survival analysis is a family of techniques that is concerned with the examination of the form and determinants of qualitative change.[4] More specifically, the focus of survival analysis is the exploration of the timing of events and the covariates which may be related to that timing. An event is defined as a shift from one mutually exclusive state to another, occurring at a specific and known point in time.

Before the advent of survival analysis, statistical treatment of event data was problematic. Consider a hypothetical example of a study comparing the recidivism rates of drug addicts who have attended a new type of treatment program with those who have received traditional treatment. Study participants are followed for 2 years. After 2 years a certain percentage of each group will have relapsed, while the rest are still clean at the

[3]The recent work of Singer, Willett, and their colleagues comes closest to providing nontechnical treatment of survival analysis (Singer & Willett, 1991; Willett & Singer, 1991a), although at a more advanced level than that of this paper. In particular, their use of survival techniques to examine the careers of teachers is exemplary (Murnane, Singer, & Willett, 1988; Murnane, Singer, Willett, Kemple, & Olsen, 1991; Willett & Singer, 1991b). They are able to show the strengths of survival analysis, especially when applied to a rich longitudinal data set.

[4]Survival analysis is also sometimes referred to in the literature as event history analysis and, less frequently, failure-time analysis. Although event history analysis is somewhat more descriptive of the techniques as they are currently used, the more common and historically rooted term is used throughout the paper.

end of the study. What is the most effective way to summarize the results to illustrate the effect of the treatment program on recidivism rates? Traditional methods would report either the percentage of each group that had relapsed during the study period, or the average time until relapse for each group. There are serious problems with each of these methods. In the former case, the proportion of the group relapsing is reported, totally ignoring the timing of the relapses. For example, each group could have a 50% relapse rate, but the control group might relapse on average 3 months sooner than the experimental group. Presenting the average time until relapse, on the other hand, has the opposite problem; it takes timing information into account, but does not indicate how much of the group has relapsed during the study.

Although both pieces of information could be presented, this would still not avoid the major problem of these methods. Both the percentage of persons relapsing and mean time to relapse are statistics that are heavily dependent on the time frame of the study. A longer period of observation almost always results in larger response proportions and higher average response times. Thus, the traditional ways of summarizing event data are dependent on the arbitrarily determined study length (Lavori, Keller, & Klerman, 1984, Singer & Willett, 1991).

Another way to restate this problem is to ask how one should handle those participants who have not relapsed by the end of the study. Every study has an endpoint that is usually arbitrary with respect to the event of interest. Therefore, there are almost always cases for which the event of interest has not occurred by the time that data collection has ended. This is one example of what is known as censored data, which is discussed below in more detail. One way to handle this problem is to count these cases as missing data. However, not only is this wasteful of participants (which is something no researcher ever wants to be!) but it results in biased estimates of average survival times. Those people with long survival times are the most likely to be censored; thus, estimates of average survival times are biased downwards if censored data are dropped. Another potential solution is to assume that the event occurs at the point of censoring. So, for the above example, everyone who has not relapsed is set as relapsing at the end of the study. Although this does not result in lost subjects, it clearly biases the survival information. Although some of the censored participants would probably relapse soon after the study is ended, many may not. Their survival time is surely longer than the length of the study, but there is not a good way to estimate that time using traditional methods.

Note that lengthening the observation period of the study (even if economically feasible) does not usually eliminate the problem of censored data. Some people may never exhibit the event of interest, for example,

many drug addicts may be able to quit "cold turkey" and never again use drugs. Also, in many naturalistic nonexperimental studies participants may enter the study at various time points. Those persons who enter near the end of the study and thus have less time to have the event occur are much more likely to be censored. So although the observation period of the study may affect the problem of censored data, altering the observation period will not be able to avoid the problem. We are still left with the need for statistical techniques that can handle censored data in an efficient and unbiased manner

One of the major strengths of survival analysis is its ability to deal with this problem of censored data. Survival analysis can handle event data without the data loss or biasing that is so problematic with other methods. Also, survival analysis can be used on many types of data from a variety of sources including prospective, retrospective, archival, and cross-sectional data (Adler & Kandel, 1983; Singer & Willett, 1991), in addition to being applicable for meta-analytic summaries of cross-study event data (Isager, Brinch, Kreiner, & Tolstrup, 1985; Lavori et al., 1984; Setze & Bond, 1985). Moreover, survival analysis lends itself to visual displays of group and subgroup survival patterns, allowing a much quicker and more intuitive understanding of the event data than quantitative summaries of event statistics. Finally, modern survival techniques offer sophisticated ways to analyze the effects of covariates on survival proportions and probabilities. For all of these reasons, survival analysis is becoming one of the most powerful tools for the analysis of change.

Modern survival analytic methods were first developed by biostatisticians interested in modeling the length of time to death of human or other living organisms (hence the term survival analysis) and engineers studying the length of time to failure of machine components (Breslow, 1974; Cox, 1972; Gail, Santner, & Brown, 1980; Kalbfleisch & Prentice, 1980; Lawless, 1982). This work in turn was partially based on the life table which is a method of summarizing event data used by actuaries and demographers since the 18th century (Pollard, Yusuf, & Pollard, 1981). More recently, survival methods have been extended to handle special cases of repeatable events and multiple types of events by sociologists studying social transitions (Blossfeld et al., 1989; Tuma & Hannan, 1984).

Although survival methods have been around for the last two decades, their use by social scientists has been sporadic. Their use seems to have caught on for only three particular research areas. First, psychiatrists have used the methods extensively to study relapse rates of persons with a variety of psychiatric conditions including anorexia nervosa (Isager et al., 1985), affective disorders (Greenhouse, Kupfer, Frank, Jarrett, & Rejman, 1987; Lavori, et al., 1984; McLeod, Kessler, & Landis, 1992), alcoholism (Hser, Anglin, & Liu, 1991; Poikolainen, 1983), suicide (Leon, Friedman, Sweeney, Brown, & Mann, 1990), and schizophrenia (Matthews, Roper, Mosher, & Menn, 1979). Second, organizational and educational psychologists have used survival analysis in more technically sophisticated ways to examine attendance and turnover in schools and at the workplace (Fichman, 1989; Harrison & Hulin, 1989; Morita, Lee, & Mowday, 1989; Murnane, Singer, Willett, Kemple, & Olson, 1991). Finally, a more disparate group of epidemiologists, sociologists, social workers, and social psychologists have been examining the prevalences and rates of more-or-less naturally occurring social phenomena such as family and relationship breakups (Felmlee, Sprecher, & Bassin, 1990; Fergusson, Horwood, & Shannon, 1984), rape prevalences (Russell & Howell, 1983), the duration of breast-feeding (Ryan & Dent, 1984), age at first intercourse (Kraft, 1991), and contraceptive discontinuation (Slonim-Nevo & Clark, 1989).

Despite the seemingly wide variety of survival analysis applications in the above list of studies, there are many large areas of psychology that would benefit by a greater use of these methods. In particular, community and ecological psychologists address many questions that could be answered with survival methods. Some examples include uncovering the factors influencing drug-use recidivism, the timing of juveniles' entrance into the justice system, prediction of violent behavior, modeling the birth and death of self-help groups, discovering the triggers to health-seeking behavior, understanding the influences of innovative program diffusion in the health care system, and discovering factors related to how long the effects last of a preventive intervention.

USING SURVIVAL ANALYSIS IN APPLIED PSYCHOLOGICAL RESEARCH

The mathematical and statistical underpinnings of survival analytic techniques can be difficult to understand. Fortunately, much of this technical background is unnecessary to be able to use, interpret, and understand survival analysis and its results. The following discussion is thus organized around three sets of practical issues a researcher typically needs to consider when performing a basic survival analytic study: (a) research design issues; (b) data preparation and organization; and (c) data analysis and interpretation.

Within each of these three general areas, many specific decisions have to be made. To provide a context for understanding how these decisions should be made, data from an existing study will be used whenever possible. The study is a large-scale evaluation of GROW, Inc. (Rappaport et al.,

1985; Roberts et al., 1991). GROW is a self-help organization for persons with a history of serious mental illness. The subset of the data presented here is used to model self-help group attendance (Luke, Roberts, & Rappaport, in press). Attendance and other data were collected over a $2^1/_2$-year period from 861 members of 15 different GROW groups in central Illinois. Summary statistics indicated that about a third of the sample attended GROW for one or two meetings and then never returned. Another third attended GROW for several months and the last third attended for longer periods of time. Although these attendance pattern summaries are interesting, they do not shed much light on more in-depth questions such as when people are most likely to stop coming to GROW meetings, what types of people are the quick dropouts, and what the factors are that lead to long-term GROW involvement. The following sections show how survival analysis can help to answer these questions.

RESEARCH DESIGN ISSUES

Minimally, to perform a survival analysis one needs only to have two pieces of information for each case: the timing of a particular event, and whether the event is censored or not.[5] However, behind this apparent simplicity lies a number of important and potentially troublesome design issues. (For other good discussions of research design and survival analysis, see Peto et al., 1976; and Singer & Willett, 1991.)

Event Timing

There are three issues that need to be considered about the timing of any particular event: event definition, event timing measurement, and the type of statistical model that will be applied to the event timing data. First is the definition of the event itself. Sometimes this is relatively straightforward, for example, the incarceration of a criminal or the pregnancy of an adolescent. (Note that these events may be difficult to measure; however, their definitions are relatively clear.) In other situations it may be more difficult to decide on a clear definition. For example, if marital breakup is the event of interest, what exactly constitutes that event? Should only formal divorces be counted, in which case the definition would be

[5]It should be noted that cases for survival analysis do not have to be persons, although they are in most of the examples discussed here. For example, the event of interest could be the dissolution of a self-help group, in which case groups are the cases, and group death is the event variable.

rather strict, or should the definition be broadened to include separations? Should other types of breakup such as death of a spouse be counted? Even when one decides on the categories to be included, there is the issue of when the event occurs. For example, assuming that only legal divorces are considered as marital breakup events, when does the divorce happen? The clearest time point may be the date of the final legal divorce decree. However, that may not be the most theoretically relevant time point. However difficult, the event must be defined so that it can occur only once for each person at a particular point in time. (The special cases of repeating events and multiple types of events are discussed later.) Essentially, the event being observed or measured should be treated like any important theoretical construct. It should be clearly conceptualized and defined so that it can be accurately and reliably measured.

In the case of the GROW attendance study, the event of interest is the point at which a person stops attending GROW meetings. Although this event is easy to define, it is much harder to operationalize and measure. The presence of an ongoing activity like group attendance is frequently easier to detect than its absence. Attendance is voluntary, and GROW members often show very different patterns of attendance. Some members attend every weekly meeting while others attend sporadically at best. (Indeed, it is this heterogeneity of patterns of attendance that makes it a substantively interesting phenomenon.)

The second issue concerning event timing has to do with whether it is measured continuously or discretely. For continuous time data, the timing of the event is measured precisely. For discrete time data, on the other hand, the event is known to occur within a particular time interval. For example, while the exact day that a person drops out of school might be known, the exact day a person became pregnant might not be as easily determined. However, pregnancy could be measured discretely, tieing the pregnancy to a particular month. Conception may have occurred at any time during that month.

In reality, event data are neither discrete nor continuous; they lie somewhere on a continuum between the extremes. There are important substantive and statistical considerations to take into account when deciding on how finely or coarsely to measure event timing. First, the time interval should be selected so as to maximize the reliability and validity of the event duration variable. Again, consider a study of pregnancy. Although conception occurs at a specific time on a specific day, in reality there is no reliable way to exactly pinpoint the time of pregnancy (at least using the methods of social scientists). However, by collapsing the time data into weekly or monthly intervals, the measure of pregnancy timing becomes more reliable.

It may be argued that by using wider time intervals important information might be lost. The issue of lost information is actually a question of validity, and is determined by theoretical considerations. For example, if the primary goal of the pregnancy study is to show how demographic characteristics such as socioeconomic and ethnic background affect the timing of first pregnancies, then measuring the point of conception exactly will be of dubious benefit. Presumably, the effects of socioeconomic and ethnic status on the likelihood of pregnancy will be felt on a monthly or yearly time scale, not daily or weekly.

The GROW data set provides another example of how reliability and validity concerns shaped the decision of the appropriate time interval. Even though raw attendance data were collected on a weekly basis, it was decided to use monthly intervals for the survival analyses. Although most GROW groups met once a week, the research staff were not always able to attend every meeting in a month. By aggregating weekly attendance data into monthly intervals, it was less likely that we would misidentify someone as a dropout when they simply attended a meeting that we missed. Simple Monte Carlo experiments using the GROW attendance data indicated that monthly aggregation produced more reliable measures of overall attendance duration. This aggregation resulted in discrete event data measured on monthly intervals. If a person stopped attending GROW, we could identify what month that happened, but not the specific day or week. After aggregation, the event history data set consisted of 27 monthly intervals.

At this point in the design process, it is useful to consider what type of statistical model will be applied to the event data. There are two broad classes of survival models: continuous time and discrete time, and each is designed for its respective type of event time data (Allison, 1984). However, just as it is not always clear exactly how discrete or continuous a particular event time data set really is, it is also not always apparent what type of statistical model is most appropriate. Not surprisingly, continuous time survival models can always be applied to event data measured continuously; and discrete time models are similarly always appropriate for event data measured discretely.

However, it is often the case that continuous time survival models are used for discretely measured event time data. For example, MacKenzie (1991), using continuous time survival models in her study of parolees, summarized recidivism data using monthly time intervals. Similarly, Adler and Kandel (1983) used yearly intervals in their investigation of risk periods for substance abuse. In both cases the time intervals made theoretical sense, but how appropriate is it to use methods designed for continuous time data on discrete data? Statisticians generally agree that when the size of the time intervals are sufficiently small relative to the rate of event occurrence, continuous time event models will produce results with little or no bias (Allison, 1982). However, as the size of the time interval increases, the size of the time-aggregation bias also increases (Heijtan, 1989; Petersen, 1991). In other words, it is safe to use continuous time models on discrete time data if the data are measured in small enough time units.

Until recently, however, there was no clear answer to the question of how small is small enough. Petersen (1991) has now provided some guidance for the optimal sizing of discrete time intervals. Using both statistical simulation and real data, he showed that when the underlying rate of event occurrence is less than .10 per time interval, the amount of time-aggregation bias is minimal, usually less than 5%. The time-aggregation bias is moderate (5 to 10%) for underlying event rates of .10 to .20 per time unit. So, for example, if it is known ahead of time that the rate of occurrence for a particular event is approximately .25 per month, then the investigator would be better off to use a weekly time interval. where the rate would be approximately .06 per week with an associated bias of less than 3%.

Despite greater calls for the use of discrete time survival models (Allison, 1982; Singer & Willett, 1991), they remain underutilized. The main reason for this is probably due to the relative difficulty of their use. Data sets must be transformed from single record per case to multiple record per case formats (Willett & Singer, 1991a). This usually requires fairly sophisticated programming to alter the record formats and create the dummy variables. Also, most of the major software packages do not have specific procedures for discrete time survival models. Instead, the researcher must use general purpose logistic regression procedures. Survival and hazard plots must be reconstructed from the produced parameter estimates. Conversely, it is much easier to use the statistical procedures for continuous time survival models. Data sets require only minimal manipulation and the major software packages all have special-purpose integrated procedures that are (for the nonstatistician) much easier to use. Because of this, the rest of this primer will primarily focus on continuous time survival models.

Event Censoring

One often does not have complete information about the event history duration for particular persons. For example, a GROW member may not have dropped out of GROW by the end of the observational study, or a GROW member may have moved out of the state during the study. In either case we do not have a measure of the event of interest (dropping out of GROW). These are called *censored* cases. Figure 1 illustrates 7 different types of censored observations (based on a figure in Yamaguchi,

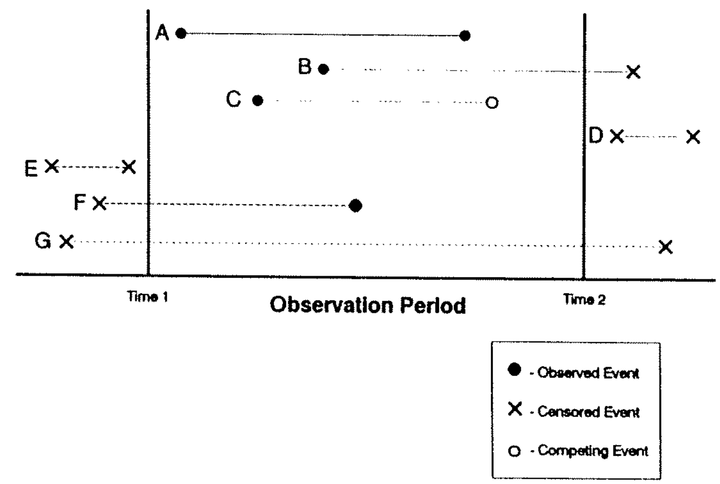


Fig. 1. Examples of censored and noncensored observation patterns.

1991). All participants are observed from Time 1 (T_1) to Time 2 (T_2), although events may occur both prior and subsequent to the observational period. The left endpoint of each participant's line marks the start of some event history, and the right endpoint marks the end of that history. The length of the line corresponds to the duration. Only Case A in Figure 1 shows an example of no censoring; the event duration begins and ends within the observation period of the study.

Event censoring may occur for one of two reasons. First, we may lack information due to the arbitrary limitations of the observation period. Events may occur either before or after the observational period, in which case event duration is not known to the researcher. In Figure 1, Cases B, D, E, F, and G exhibit this type of censoring. A second type of censoring occurs when a participant is lost to observation due to an event other than the event of interest.[6] For example, in a study of school dropouts a participant might be lost because her family moved to a different city. This person should not be considered a dropout, however her exact event du-

[6]This is sometimes known as competing risks, because other events can occur that would exclude the event of interest from occurring. Competing risks can be modeled in their own right, and are discussed below.

ration is not known because of the move. Case C in Figure 1 is an example of this type of censoring.

Right censoring occurs when a participant has not exhibited the event of interest by the time observation ends, or the event duration has ended for reasons other than the event of interest. Cases B, C, D, and G in Figure 1 are examples of right-censored observations. One of the major advantages of survival analysis is its ability to handle right censored data. As long as censoring is unrelated to event occurrence, survival analysis can produce unbiased estimates of survival time summary statistics (Greenhouse, Stangl, & Bromberg, 1989). What this means is that the pattern of censoring cannot give us information about the likelihood of the event of interest. For example, in a study of job attendance, moving to a different state may be considered independent of the likelihood of staying home from work. On the other hand, death due to illness may be related to the event of interest—ill people are more likely to miss work for health reasons. Most of the major software packages allow the examination of censoring patterns in order to test for independence.

Although right censoring is well covered in the survival analysis literature, there is another type of censoring that is both more problematic and not often discussed. *Left censoring* occurs whenever duration information is missing because of events that happen before the start of the observational period. In Figure 1, Cases E, F, and G exhibit left censoring. One of the reasons that left censoring is not discussed as much in the literature may be due to the fact that left censoring never occurs in experimental designs where subject participation is controlled by the researcher. In this case, all participants enter into the study at the same time, and event durations cannot start before the observational period starts. However, left censoring may often occur in other types of designs, including quasi-experimental, retrospective, and archival studies. In general, survival analysis cannot adequately handle left-censored data. Therefore, it is imperative that left-censored cases not be included in a survival analysis.

For an example, consider the GROW attendance study. Many GROW members had been attending meetings before the research staff started observing the groups. Although the date these people stopped attending GROW could be observed, their starting date happened some time before the study began. This type of observation is represented by Case F in Figure 1. Even if we could know when they started (from self-report or GROW records), survival analysis would not be able to provide unbiased survival estimates. This is because by including all those persons who are already attending GROW at the start of the study, we are more likely to be including data from those persons who have attended GROW for relatively longer periods of time. Persons who attend GROW for shorter periods of time are more likely to have started and stopped attending before the beginning of the observational period. An example of this is Case E in the figure. If left-censored cases are included, estimates of average survival times will be biased upward. Thus, for nonexperimental studies it is extremely important to make sure that no left-censored cases are included in the analyses. For the GROW data this is done by only including those GROW members who started attending GROW after the observational period started. (For a more detailed discussion of censoring and its implications see Little & Rubin, 1987; or Yamaguchi, 1991.)

DATA PREPARATION AND ORGANIZATION

Once a survival analysis study has been designed and the data collected, there are some data management steps that need to be done before the analyses can be performed. Many traditional multivariate methods can be performed on unmodified rectangular data sets. However, it is almost always necessary to translate from a raw data set into one that is appropriate for survival analysis. In particular the dependent variable, which is the event being modeled, is actually made up of two discrete pieces of information: the event time and the event status. The event time is a nonnegative value that indicates the participant's time to event or response. Event time may be measured at any interval scaled level that makes sense for the particular study (e.g., seconds, days, months, years). Event status is a categorical variable that indicates if the event has occurred, or if it is censored. (Several computer programs allow multiple types of censoring codes to indicate, for example, if a subject is lost to follow-up, has not responded at end of study, etc.)

The first step in data organization is to calculate the event time variable. Sometimes this is relatively straightforward, as it would be in a study of school dropouts where the total number of months in school is directly recorded. In this case number of months in school is equivalent to number of months until dropout. If a person has not dropped out by the end of the study, the total number of months in school at the end of the study would be used in the event time variable. Other study designs may require more complicated data management. For example, in nonexperimental studies where participants can enter observation at varying times, the time event variable often must be calculated as the difference between a time-entered variable and a time-exited variable. In the GROW study time until dropout was calculated by subtracting the date the participant first attended

Table I. Comparison of Raw and Translated Data Sets

	Raw data set			Tanslated data set		
Case	Start date	End date	Gender	End months	Status	Gender
1	05/06/84	07/28/84	1	3	1	1
2	05/07/84		2	27	0	2
3	06/02/84	12/05/85	2	19	1	2
4	11/01/85	03/23/86	1	5	1	1
5	05/01/86		2	3	0	2

a GROW meeting from the date of the last meeting attended. If a person was still attending at the end of the study, the last date of observation was used to calculate the time event variable.

To see this more clearly, consider Table I. The left side of the table presents the raw data for five example cases from the GROW study. Each person has a start date, and those persons who stop attending during the observational period have an end date. If a person was still attending at the end of the study, they have no end date (indicated by a blank). (Remember, these cases are right-censored.) The right side of the table presents the translated data that can be used by the survival analysis. The event time variable "End Months" is simply calculated as the month after joining GROW that the person stopped attending. For example, Person 1 stopped attending in the third month after he or she started attending. Calculation of the event time variable is slightly more complicated for right-censored cases. Here, the end date is assumed to be equal to the date of the end of the study or observational period. For this example, the end of the study was July 31, 1989, 27 months after the start in May of 1987. Therefore, for Person 2 who attended the entire duration of the study, the event time variable is 27.

The event status variable indicates how the event time variable is to be interpreted by the survival analysis program. In this case, the status variable is coded 1 for a response, and 0 for a censored observation. A response means that the event of interest has occurred, in this case a person was observed who stopped attending GROW meetings. Thus in Table I Person 1 who was observed to drop out of GROW after 3 months has a status variable of 1, whereas person 5 who had been in GROW for 3 months at the time the study ended has the same event time value (3), but the event status variable would be coded 0 to indicate a censored observation.

Covariates are included in the data set just as they would be for a typical multivariate analysis. In Table I, for example, each case has a value for the variable Gender, and this value is the same in both the raw and translated data sets. Of course, any number of covariates may be included in the data set.

DATA ANALYSIS AND INTERPRETATION

Describing the Overall Survival and Hazard Functions

After the event of interest has been defined, the event durations have been observed, and the data have been translated, the first analytic step is often to examine the estimates of the overall survival and hazard functions. With this information you can discover some basic facts about the general pattern of survival times for the sample as a whole. In effect these estimates provide summary descriptive statistics for the event history data.

There are two commonly used nonparametric methods for estimating the overall survival functions. The *life table* or actuarial method groups the event durations into time intervals (e.g., weeks, months, years), summarizing the results in a life table. *The product-limit* or Kaplan-Meier estimate, on the other hand, uses ordered observations rather than grouped data (Kaplan & Meier, 1958). Here, all cases are ordered from shortest to longest event duration. The product-limit estimate has the advantage of producing results that are not dependent on the size of any time intervals, and is especially useful for small sample sizes (Slonim-Nevo & Clark, 1989). The product-limit and life table approaches yield identical estimates when there is no censoring and no ties (each time interval contains at most one observation) (Dixon, 1990). If you have discrete event time data and/or a large number of event duration ties, you should not use the product-limit estimate. However, if you have continuous event time data you can use either method.

Table II presents the life table for the GROW attendance data. (Unless otherwise noted, all statistical analyses were done using BMDP 1990 Version, programs 1L and 2L. All graphs were produced based on these data using Harvard Graphics for Windows.) A life table is a way of summarizing event history data. It is usually too detailed to appear in published journal articles, but it is important to understand how it is organized because the concepts of survival and hazard functions are based on information presented in the table.

Each row in the table represents one time interval, in this case a month. The entries in the *N* column indicate the number of cases that have survived (are still attending GROW) to the beginning of the current interval. So, for example, in the first month after joining GROW (the second line in the table) all 644 persons are still attending, since by definition you

Table II. Life Table for Overall GROW Attendance

Interval start months	N	No. with-drawing (censored)	No. exposed to risk	No. terminal events	Pro-portion termi-nating	Pro-portion surviving	Cumulative proportion surviving	Hazard
0	644	0	644.0	0	.00	1.00	1.00	0.00
1	644	7	640.5	249	.39	.61	1.00	.48
2	388	13	381.5	81	.21	.79	.61	.24
3	294	5	291.5	35	.12	.88	.48	.13
4	254	8	250.0	23	.09	.91	.42	.10
5	223	5	220.5	24	.11	.89	.37	.12
6	194	6	191.0	14	.07	.93	.34	.08
7	174	3	172.5	14	.08	.92	.32	.08
8	157	5	154.5	8	.05	.95	.29	.05
9	144	0	144.0	12	.08	.92	.28	.09
10	132	5	129.5	7	.05	.95	.25	.06
11	120	2	119.0	10	.08	.92	.24	.09
12	108	2	107.0	10	.09	.91	.22	.10
13	96	5	93.5	10	.11	.89	.20	.11
14	81	2	80.0	8	.10	.90	.18	.10
15	71	3	69.5	1	.01	.99	.16	.01
16	67	3	65.5	12	.18	.82	.16	.20
17	52	0	52.0	5	.10	.90	.13	.10
18	47	1	46.5	1	.02	.98	.12	.02
19	45	3	43.5	2	.05	.95	.11	.05
20	40	6	37.0	7	.19	.81	.11	.21
21	27	0	27.0	4	.15	.85	.09	.16
22	23	2	22.0	3	.14	.86	.07	.15
23	18	5	15.5	2	.13	.87	.06	.14
24	11	1	10.5	3	.28	.71	.06	.33
25	7	2	6.0	1	.17	.83•	.04	.18
26	4	2	3.0	0	.00	1.00	.03	.00
27	2	2	1.0	0	.00	1.00	.03	.00

cannot drop out before you have started attending.[7] The Number Withdrawing entries are the number of cases that are censored in that particular interval. For the first month after joining GROW, we see that seven persons are censored during this interval. These seven joined GROW very near the end of the study. They are still attending GROW at the end of the observation period and we do not know when, if ever, they stopped attending GROW. By the above definition they are right censored. The Number Exposed to Risk is defined as the number of cases entering the

[7]There is an important difference between the observation period of the study and the attendance period as presented in Table II. The 644 people who enter the analysis in the first month did not all start attending GROW in the first month of the study. Rather, they started sometime within the 27-month observation period. The first column in the table, Interval Start, always refers to the number of months a person has survived to, regardless of the month of the study.

interval minus one half of those censored during the interval. By only including half of the censored group, the survival time statistics are adjusted to provide unbiased estimates. This number is calculated this way because censored observations have not been observed for the entire interval. The assumption is that they have been observed, on average, for half of the length of the interval. This assumption is one of the reasons that survival analysis can use censored data and produce unbiased survival estimates. The number exposed to risk is 644 – 0.5*7 = 640.5 The Number of Terminal Events entry is the number of cases for whom the terminal event (stopping attending) has occurred in that interval. In the first month 249 people stop attending GROW. The Proportion Terminating is simply the number of terminal events divided by the number exposed to risk. The Proportion Surviving is one minus the proportion terminating. In the first month then, the proportion terminating is 249/640.5 = 0.39, and the proportion surviving is 1 – .39 = .61.

The last two columns in the life table actually provide the estimates of the survival and hazard functions. The Cumulative Proportion Surviving is an estimate of the probability of surviving to the beginning of the *i*th interval, or longer, without occurrence of the event. It is computed as the product of the proportion surviving this interval and the proportion surviving all previous intervals. For the GROW data, then, there is a 100% chance of surviving to the beginning of the 1st month. There is a 61% chance of surviving to the 2nd month (1.00*0.61), and a 48% chance of surviving to the 3rd month (.61*.79). [Note that the estimated probability is not the same as the raw proportion of participants surviving. For example, the raw proportion surviving to the 2nd month is .602 (388/644). Only when there is no censorship is the estimate equal to the simple proportion. Again, this is because the survival estimate takes censored cases into account.]

Figure 2 presents the graph of the estimated survival function from the life table. The horizontal axis represents time, and the vertical axis shows the proportion of cases still surviving. The survival curve is a monotonically nonincreasing function, meaning that it can only stay the same or get smaller over time. To interpret the survival curve, you look at a number of things. First, the slope represents the length of survival time: a shallow curve means longer survival times while a steep curve indicates shorter survival times. Sudden changes in the slope over time indicate shifts in the probability of survival. The survival graph in Figure 2 indicates that a large number of the sample stop attending GROW very quickly, as shown by the steep curve at the beginning of the graph. After the first few months the curve flattens out, suggesting that the rate of dropouts slows down. The median survival time can be found by drawing a horizontal line from the .5 point on the Y axis until it intersects the curve. For the GROW data the median survival time is 2.86 months (this is also printed out by the computer program), indicating that half of the sample has stopped attending GROW meetings in less than 3 months. The median survival time is a more useful summary statistic than the average survival time due to the existence of right-censored cases. For censored cases, we do not know exactly when the event will occur, if ever. Therefore, average survival time cannot be accurately calculated.

The last column in Table II presents the estimated hazard rate, which for continuous time models refers to the instantaneous rate of event occurrence, given that the event has not occurred up to that point. This is also known as the event *risk*.[8] So whereas the estimated survival rate tells us how many people have survived up to some time point, the estimated hazard rate tells us the event rate for any particular person within a specific time interval, assuming they have survived to that time interval. In this case the time interval is 1 month. We see, for example, that the risk of dropping

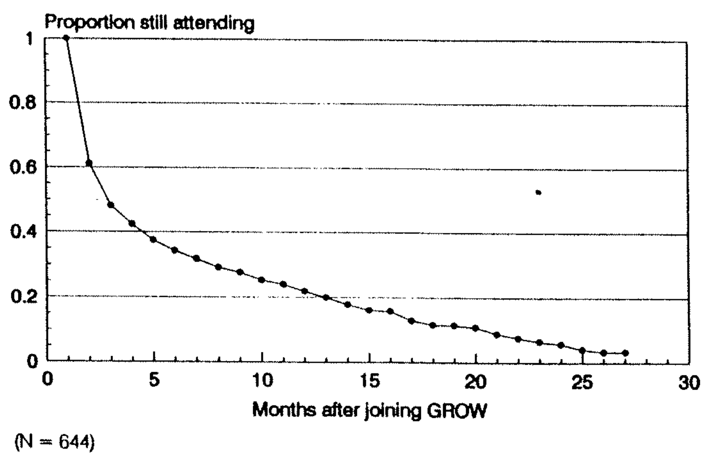


Fig. 2. Estimated overall survival function for GROW data.

[8]For discrete time models, the hazard rate is always calculated for each time interval, and can never exceed 100%. In this case the hazard rate is appropriately interpreted as the probability of occurrence in that particular time interval, given that the event has not occurred in a previous interval. However, when event times are measured continuously, there is no time interval. (More accurately, the time interval i is infinitely small.) In this situation the hazard rate can actually exceed 100%. The greater the risk, the more likely that the event will occur at that instant in time.

out of GROW is relatively high in the second month after joining (.24), provided that one has not dropped out during the first month. Again, the estimated hazard rate is not equivalent to the simple proportion of people terminating to number of persons exposed to risk within a specific time interval.[9]

Figure 3 presents the graph of the estimated hazard function for the GROW data. Each point on the graph represents the hazard rate for a particular month interval. It shows that the risk of dropping out of GROW is high for the first 2 months after joining, after which it levels off at approximately .08 to .10 for the next year. Around 15 to 18 months the dropout risk starts rising again. This suggests that GROW members are most likely to drop out very early after joining. Also, if they stay in for at least 1.5 years, they become more likely to drop out after that time.

Examining the overall hazard function graph is a quick way to see how the rate of event occurrence changes over time. Used this way, the

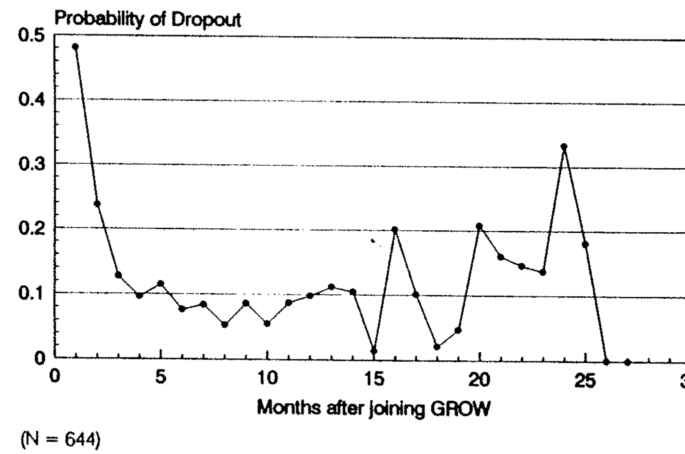


Fig. 3. Estimated overall hazard function for GROW data.

[9]Unlike with the estimated survival function, it is more difficult to see how the numbers for the estimated hazard rate are calculated. The equation for the estimated hazard rate is:

$$h_i = 2q_i/w_i(1 + p_i),$$

where q_i is the proportion terminating for some timepoint i, p_i is the proportion surviving, and w_i is the width of the time interval (SPSS Inc., 1991). This equation may not be intuitively obvious, but it has been found to be an unbiased estimate of the underlying population hazard rate (Gehan, 1975). So, for the first month after joining GROW, the hazard rate is (2*.39)/(1*(1 + .61)) = .48.

hazard function is much more substantively useful than the survival function because the latter can only decrease over time, while the former can go up or down depending on the underlying event probabilities. In fact, analyzing the shape of the hazard function is a primary task in survival analysis, as we shall see later.

Modeling the Effects of Covariates

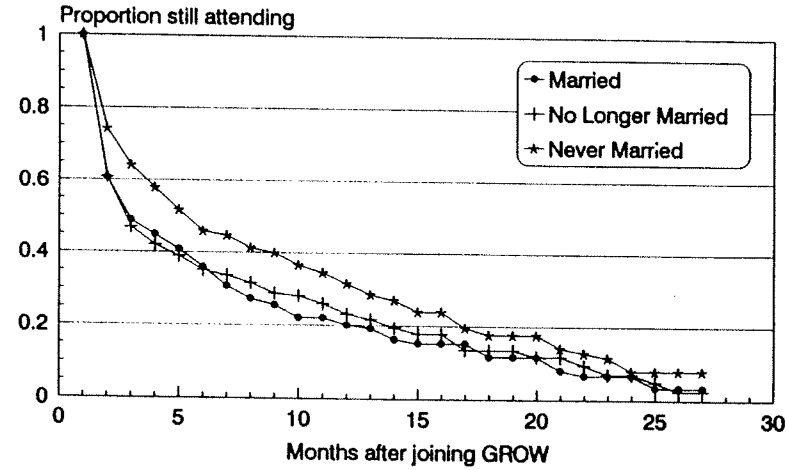
A description of the overall survival and hazard functions can provide interesting information about the overall patterns of events over time.. At the same time, however, these general curves based on the entire sample can mask theoretically important heterogeneity. When examining the effects of covariates on event durations, investigators can either examine single predictors one by one, or examine sets of multiple predictors.

Single Covariate (Subgroup) Analyses

Persons belonging to different subgroups defined by a single categorical covariate may exhibit quite different survival patterns. For example, girls in general may be less likely to drop out of school than boys. If this were the case, then girls would have a higher survival curve relative to the boys', indicating a higher proportion of girls staying in school for any particular point in time. Another typical situation calling for subgroup analyses is when survival times for an experimental group are compared to those of a control group. Subgroup analyses are probably the most common example of survival analysis found in the social science literature (e.g., Fisher & Anglin, 1987; Greenhouse et al., 1989; Isager et al., 1985; Ryan & Dent, 1984)

Subgroup analyses within survival analysis are quite straightforward. Essentially, a product-limit or life table is produced for each subgroup separately. Survival and hazard curves can then be plotted for each subgroup simultaneously, aiding in the interpretation of any subgroup differences. This can be done automatically by most of the statistics packages simply by specifying which categorical variable should be used. For example, specifying gender as a subgroup covariate would produce two sets of survival estimates: one for women and one for men.

Figure 4 shows the survival curves for an example subgroup analysis of the GROW attendance data. This analysis is based on a single categorical predictor: marital status. The graph indicates that persons who have never been married are more likely to stay involved in GROW, while those who are or were at some time married drop out of GROW at a relatively faster rate. Almost 40% of the persons who are or have been married drop



(N = 467; Mantel-Cox = 7.46; p = 0.024)

Fig. 4. Subgroup analysis: Duration of attendance by marital status.

out of GROW after only 1 month, compared to only 25% of the never married group. Also, the median duration of attendance for the married or previously married groups is slightly less then 3 months compared to just over 5 months for the never married group.

There are three common nonparametric statistics used to test the null hypothesis that there are no differences between subgroup survival functions based on single categorical predictors. Use of these statistics allows a more formal test of group differences than mere visual inspection would allow. Each of these statistics are log-rank tests that are variations of Wilcoxon tests (Hays, 1963). The most commonly used statistic is the Mantel-Cox, or simple log-rank test (Mantel, 1966), which applies equal weights to every time period. The Breslow statistic gives more weight to earlier observations, thus it is less sensitive to later events than the Mantel-Cox test. The Tarone-Ware statistic uses an intermediate weighting scheme and is a compromise between the Mantel-Cox and Breslow statistics (Dixon, 1990; Tarone & Ware, 1977). Unless there is a strong theoretical reason to employ a weighting scheme, the Mantel-Cox test should be examined first. For example, if you believe that the important effects of an intervention program will occur in the first few months after program participation, you might want to use the Breslow statistics to emphasize any early subgroup differences. This might be especially useful if you have a small sample, which would normally decrease your chances for discovering important subgroup differences.

A detailed discussion of sample size and statistical power relating to survival analysis is beyond the scope of this paper. The minimum required sample size depends on the desired effect size, the number of groups being compared, and the length of the observation period. Large numbers of censored cases reduce the power of the tests. Also, the tail of the survival curve often contains very few cases. This can sometimes make it more difficult to detect group differences. If you expect to see most of the interesting subgroup differences relatively early, then you can "cut off" the survival curve at the point where n becomes too small. Another possibility is to compare the Breslow statistic to the Mantel-Cox. If there is a significant group difference using the Breslow, but not the Mantel-Cox, then there is evidence that the largest group differences are occurring earlier on in time. Singer and Willett (1991) provide an excellent overview of this literature.

Multiple Covariate Analyses

Subgroup analyses, despite their usefulness, have two major limitations. First, by definition they are restricted to consideration of categorical variables. Continuous variables such as age can only be utilized after they have been transformed into categorical variables using rational or empirical criteria like median splits. This transformation results, of course, in a substantial loss of potentially important information. The second, and more serious problem is that only one covariate (categorical variable) can be analyzed at a time. Just as moving from simple regression to multiple regression allows more sophisticated modeling, the ability to handle multiple covariates greatly enhances the usefulness of survival methods.

In 1972, Cox presented a method that addressed both of these problems. The revolutionary article "Regression Analysis and Life Tables" described what is now known as the proportional hazards model. Not only does this method allow for multiple continuous and categorical explanatory variables, but these covariates can also be time-varying. Thus in a study of job transitions, socioeconomic status (categorical), age at time of first employment (continuous time-invariant), and salary at each yearly review (continuous time-varying) are all explanatory variables that can be simultaneously modeled using Cox's proportional hazards technique.

A major strength of the proportional hazards model is that it makes no assumptions about the form of the hazard function over time; it can take any shape whatsoever. By ignoring the exact shape of the underlying hazard

function, the focus shifts to how explanatory variables may or may not be related to the hazard function. Because the proportional hazards technique does not model the shape of the hazard profile, but it does model the effects of covariates on the hazard function, this is known as a semiparametric model. If an explanatory variable is related to the hazard function, then changes in this variable are associated with changes in the overall level of the hazard function in a linear fashion. That is, if an explanatory variable is related to hazard, this relationship will show up in a shift up or down of the hazard curve. For example, if men are likely to get arrested sooner than women after being released from prison, then a proportional hazards test of this hypothesis would show a hazard curve for men that has been shifted upward relative to that for women. The major assumption of the proportional hazards model is that for any two cases the ratio of their hazards is a constant across all time points (i.e., proportional) (Cox, 1972).

Table III shows the results of a proportional hazards model testing the relationship between a set of individual level variables and the probability of dropping out of GROW. Although somewhat cryptic, the material presented in this table provides a wealth of important information about the relationship between the covariates and the hazard function.

The first column of the table lists each of the covariates. In the next two columns the parameter estimate (coefficient) and its associated standard error are listed for each variable. The sign of each coefficient indicates the direction of the effect (if any) of the covariate on the hazard function. A positive coefficient increases the hazard function (shifting it up), thus decreasing survival time. So, persons with higher psychological functioning have a higher hazard rate, resulting in dropping out of GROW more quickly.

To determine which covariates are significantly related to the hazard rate, the ratios of the estimates to their standard errors can be examined. These ratios, listed in the fourth column, are z scores and are interpreted like t-tests in ordinary multiple regression. Thus, any ratio with an absolute

Table III. Proportional Hazards Model of Individual Level Variables[a]

Variable	Coefficient	Standard error	Coefficient/SE	Exp(coeff.)
Gender	0.1408	0.1473	0.9557	1.1512
Age	−0.0126	0.0059	−2.1133[b]	0.9875
Marital status	−0.1978	0.0931	−2.1242[b]	0.8206
Educational level	−0.1265	0.0700	−1.8067	0.8812
Psychological functioning	0.2077	0.0675	3.076[c]	1.2309

[a]Global χ^2 (5) = 20.34, p = .0011.
[b]p < .05.
[c]p < .01.

value over 1.97 indicates a significant covariate at the .05 alpha level. Table III shows that age, marital status, and psychological functioning are significant, with educational level showing a trend level effect.

The last column provides the most useful information for interpreting the effects of the covariates. The raw coefficient is difficult to interpret directly, because the Cox model is actually estimating the natural logarithm of the hazard rate. Interpretation becomes simpler if we exponentiate (take the antilog of) the coefficient. [If d is the coefficient, compute exp(d); i.e., raise e to the d^{th} power.] Now for each unit increase in the explanatory variable, the hazard rate is multiplied by the exponentiated coefficient. For example, a unit increase in age results in the hazard rate being multiplied by 0.9875. Thus, the hazard rate is reduced by 1.25% for every year added to age. (1.0 — .9875 = .0125.) This type of interpretation holds for non-continuous variables as well: Moving from one lower level of psychological functioning to the next higher level results in a 23% increase in the hazard rate ($h(t)$*1.2309). Taken together, these results show that GROW members who are young, married, and high functioning are more likely to drop out of GROW early than those who are older, have never been married, and are low functioning.

? 1/.98

The proportional hazards model is an extremely flexible survival analysis technique, and it provides numerous ways to examine the data. Most Cox model algorithms allow for stepwise as well as manually specified covariate entrance and removal. (Although stepwise regression is possible, one should use this technique with extreme caution, especially for hypothesis testing. See Hocking, 1983; Lovell, 1983; Wilkinson, 1979.)

Another useful technique is to compare survival curves based on user-specified covariate patterns. These are also sometimes called covariate vectors. For example, consider a study examining the effects of gender and political orientation on political involvement. Gender is coded as 0 = female and 1 = male, and Politics is coded as 0 = liberal and 1 = conservative. Separate survival curves could be estimated for liberal males (1,0), conservative females (0,1) and so on. Most of the major computer packages allow any arbitrary covariate pattern to be used to estimate the survival function (see Appendix).

Figure 5 compares two extreme covariate pattern survival curves for the GROW data with the complete sample (baseline) survival curve. The first covariate pattern represents the best-case GROW member: older, never married, college degree or better and low functioning. These are the people who are most likely to stay in GROW the longest. The other curve is the estimated survival function for the worst-case members: younger, married, high functioning, and less than high school educated. The differences in the survival curves are striking. Within the first half year after joining GROW, essentially everyone who is in the latter group has dropped out, while the former group is likely to have over 35% still involved with GROW 2 years after joining. After identifying significant predictor variables via the Cox proportional hazards model, various covariate vectors can be constructed to help find out what types of people have what sort of survival patterns. This ability to jump back and forth between exploratory and hypothesis-testing modes of enquiry makes survival techniques intriguing and engaging tools.

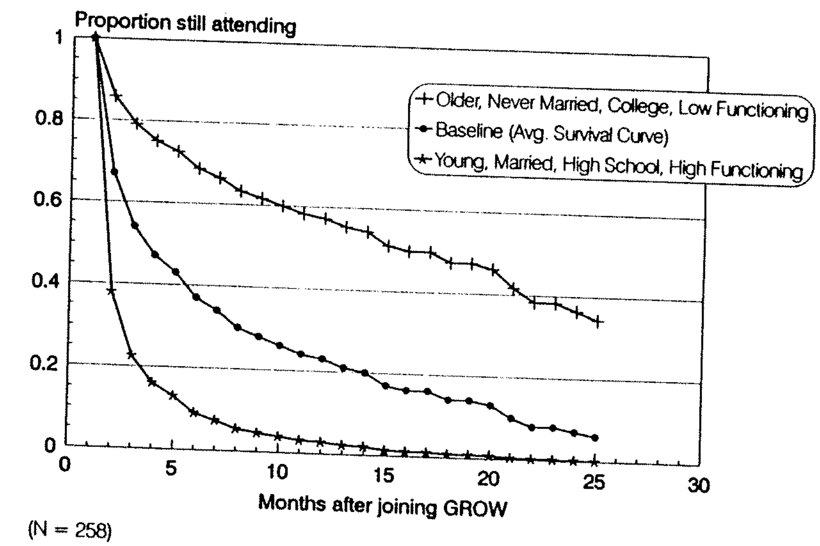


Fig. 5. Comparison of baseline and covariate vector survival curves: Duration of attendance.

Modeling Hazard Function Form

Although the above techniques are very powerful, they tend to impose a reductionist viewpoint on the analyses. That is, the survival analysis is interesting to the extent that individual variables or subgroups are found that are related to the overall survival curve or hazard function. The researcher never knows if important variables have been overlooked, or important subgroups missed. Use of the Cox semiparametric model is similar to an ANOVA or multiple regression approach—the focus is on the covariates rather than on the structural form of the overall hazard function.

However, another advantage of survival analysis is that it is not limited to an inspection of covariates. The overall form (i.e., shape) of the hazard curve over time can be examined and modeled. This can reveal important aspects of the event under consideration. For example, are children more or less likely to be absent as the school year progresses? Do computers break down randomly, or is it more likely within the first 72 hours, as some manufacturers suggest? Does the probability of drug use recidivism increase over time? These are all questions that can partially be answered by modeling the form of the hazard curves for particular events. The focus is on the probability of an event within a particular time period, not which variables may be related to that probability.

To model the form of the hazard function over time, a particular parametric model can be chosen. Choice of the appropriate model should be based on theory, although empirical results and statistical convenience usually come into play. Unlike the Cox semiparametric model, which intentionally ignores the form of the underlying hazard function, all fully parametric models estimate parameters for the hazard function. In general, a parametric model estimates two parameters describing the function. A location parameter indicates the average height of the hazard function, while a shape parameter shows shape or slope of the function.

All hazard curves (and therefore, the functions that model those curves) can be categorized into one of three classes:

1. The hazard is constant over time. This means that the probability of the event occurring does not change as time passes. For example, in physics the probability of radioactive decay is constant over time. A graph of this curve would show a single horizontal line. The exponential model[10] describes a flat, nonincreasing or nondecreasing hazard rate. Although the exponential model is easier to understand than other parametric models, its use is limited for social scientists. Most events of interest to social scientists have hazard rates that are dependent on time. Other parametric models must be considered when a nonconstant hazard rate is assumed.

2. The hazard either increases *or* decreases over time. As time passes, the probability of event occurrence gets either larger or smaller. This type of hazard is quite common. For example, in organizational psychology a theory of work attendance that views absences as a product of accumulated stress would hypothesize an increasing hazard rate over time. As stress builds up, a person is more likely to stay home from work. On the other hand, others have theorized that attendance is a habit. The more one stays

[10]Calling this the exponential model can be confusing, given that the model describes a flat, linear probability curve. Although the hazard is constant, the distribution of time until event occurrence is exponential.

at work, the more that the person is likely to continue staying at work. This results in a decreasing hazard over time (Harrison & Hulin, 1989). The two most common parametric models used to describe increasing or decreasing hazard rates are the Gompertz and Weibull models. Although they do differ from each other, these differences are often of more statistical than theoretical importance. (See Blossfeld et al., 1989, or Steinberg & Colla, 1988, for more detailed discussions of these parametric models.) The important similarity is that they both include parameters for the shape of the hazard function. The shape of the parameter is generally greater than one for an increasing function, and less than one for a decreasing function.

3. The hazard increases *and* decreases over time. In this case the hazard rate increases over some particular time span, and decreases over another. The most common example of this form is probability of death. The probability of death is fairly high early on in life, decreases as a person goes through adolescence and adulthood, and then increases in the later years of life. Unfortunately, there is no commonly used parametric model that describes U-shaped hazard curves such as this. On the other hand, two parametric models are available if an inverted U-shaped curve is proposed (low hazard, high hazard, then low hazard again): the log-normal and log-logistic.

As suggested above, the technical sophistication of these models often outpaces social theory. That is, our hypotheses and data usually are not precise enough to distinguish between the Gompertz and Weibull models, or the log-normal and log-logistic models (Allison, 1984).

How then, should these models be used in everyday applied research? Often, the most important task is to determine if the hazard rate is dependent on time in any simple fashion. This can be established by a few simple steps that are relatively straightforward and included in any survival analysis package that supports parametric modeling.

The first step should be to plot the hazard rate and examine it visually. Consider Figure 6, for example, which shows the hazard rate curve for a group of GROW members who attended at least 3 GROW meetings.[11] The plot indicates that for the first $1^1/_2$ years after joining GROW, the risk of dropping out remains constant at about .08 to .10. After 18 months, the

[11]A subgroup of the complete GROW attendance data set is being used here for two reasons. First, examination of the hazard rate for the entire sample suggests a U-shaped function (cf. Figure 3). As discussed above, it is not possible to adequately model this type of hazard curve. Second, other results suggest that the GROW sample is made up of two distinguishable populations: a group of persons who attend GROW once or twice and never come back, and those who attend semiregularly for months or years (Luke & Roberts, 1992). To validly model hazard rates it is important not to have more than one homogeneous population included in the analysis. See the discussion of unobserved heterogeneity.

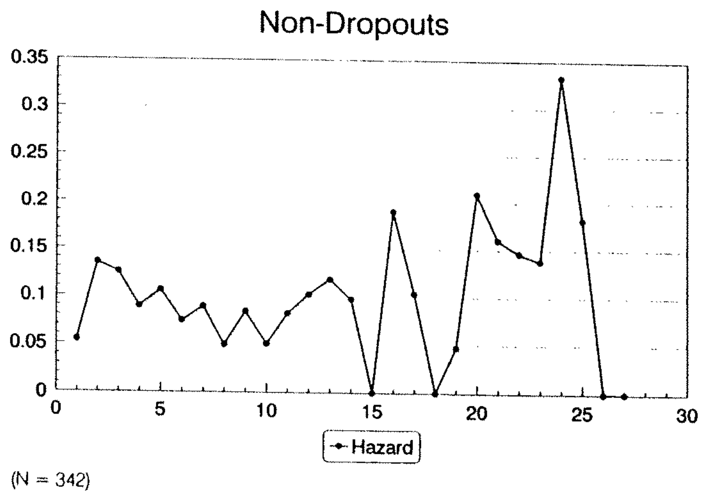


Fig. 6. Estimated hazard rate for non-dropouts.

hazard suddenly rises to somewhere between .15 and .20. Thus, this graph suggests that there is a long-term positive relationship between the hazard of dropping out of GROW, and the length of time of GROW membership.

The next step is to test more formally the hypothesis that the GROW attendance hazard rate does increase over time. There are a variety of ways to do this. Lawless (1982) shows how to use residual plots to test for departures from exponentiality. (Remember that a constant hazard rate is described by an exponential function.) Another, more straightforward approach is to fit both the exponential model and one of the increasing/decreasing models such as the Weibull, and then test the relative fit of the two models (Allison, 1984). This works because the exponential model is a special case of the Weibull, Gompertz, and other models. The log-likelihoods of each model can be compared. (The log-likelihood is a goodness of fit statistic that derives from the maximum-likelihood estimation procedure in parametric models.) If the exponential model provides a poor fit to the data (i.e., the hazard rate is not constant over time), then its log-likelihood will differ significantly from that of the alternative model. This is similar to testing for a significant difference in change of R^2 in multiple regression. (See Kennedy, 1983, for a good introduction using log-likelihoods in this manner, and see Allison, 1984, for an application of this method to survival analysis.)

Table IV shows the results of fitting both the exponential and Weibull models to the GROW data. (This table is based on results obtained from

Table IV. Exponential and Weibull Parametric Models of GROW Data

Covariate	Exponential model				Weibull Model			
	Coefficient	*SE*	Coefficient/*SE*	Exp(coeff.)	Coefficient	*SE*	Coefficient/*SE*	Exp(coeff.)
Gender	−0.13	0.18	−0.71	0.88	−0.12	0.13	−0.88	0.89
Marital status	0.17	0.11	1.48	1.18	0.14	0.08	1.70	1.15
Education	0.14	0.08	1.77	1.15	0.13	0.06	2.18	1.14
Age	0.44	0.17	2.55	1.55	0.39	0.13	2.95	1.48
Location	3.78	2.04	1.85	43.8	4.41	1.81	2.43	82.3
Shape	1.00				1.21	0.09	13.25	3.35
	Log-Likelihod: −656.23				Log-Likelihood: −651.29			

the SURVIVAL module of Systat.) Important covariates should be included in both models, and the covariate parameters are interpreted just as they are using the Cox model. Parametric models also estimate coefficients for the form (i.e., location and shape) of the underlying hazard model. The location parameter is usually not of any substantive interest. The shape parameter for the exponential model is constrained to 1.00, indicating that the hazard rate is assumed to be constant over time. For the Weibull model, the estimated shape parameter for these data is 1.21, indicating a possible increase in the hazard rate over time. (A value of less than 1 would indicate a possible decreasing hazard rate.)

The Weibull model contains all the variables included in the exponential model; in addition it includes the shape parameter. Any difference in the overall log-likelihoods of the two models is due entirely to this additional parameter. To test this, take the positive difference of the two likelihoods, and multiply it by two. This test value is distributed as a chi-square variable with degrees of freedom equal to the difference of the degrees of freedom of the two models (Allison, 1984). For the GROW data the difference between the two likelihoods is $656.23 - 651.29 = 4.94$. Twice this is 9.88. This value is compared to the chi-square critical value of 3.84 ($df = 1$, $\alpha = 0.05$). This significant result indicates that the exponential model does not adequately fit the data. Thus, there is evidence to suggest that the hazard rate for dropping out of GROW increases over time.

Estimated?

Although this process sounds complicated, it is actually much easier to do than to describe. By comparing the two log-likelihoods (which are similar to χ^2 statistics), we are simply finding out if there is evidence to suggest that the hazard rate changes over time. Because we found that there was a significant difference between the two models, we knew that the hazard rate must change over time, since the only difference in the two models was the inclusion of a shape parameter in the Weibull model. Furthermore, because the value of this shape parameter was greater than 1 (1.21), we also know that the hazard function increases over time. Referring to Figure 6 we can see that this test has confirmed what we already suspected from visual inspection of the estimated hazard function.

There is one problem common in survival analysis data that may lead to erroneous interpretations of hazard rates. If a sample actually consists of distinguishable subpopulations, each of which has a different but constant hazard rate, the hazard curve may appear to decrease over time even though every person in the sample has a constant hazard. This is known as the problem of *population heterogeneity* or *unobserved heterogeneity* (Blossfeld et al., 1989). Figures 7 and 8 illustrate this problem using a fictitious example. We have three subgroups of 100 individuals. Each subgroup has a constant hazard rate, but the level of hazard is different for

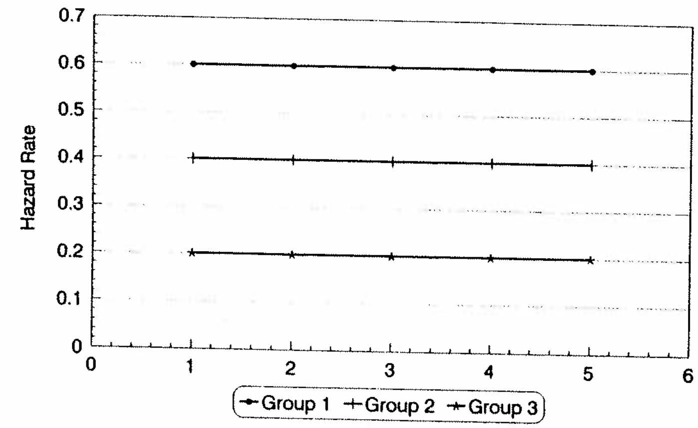


Fig. 7. Example of unobserved heterogeneity with three subgroups.

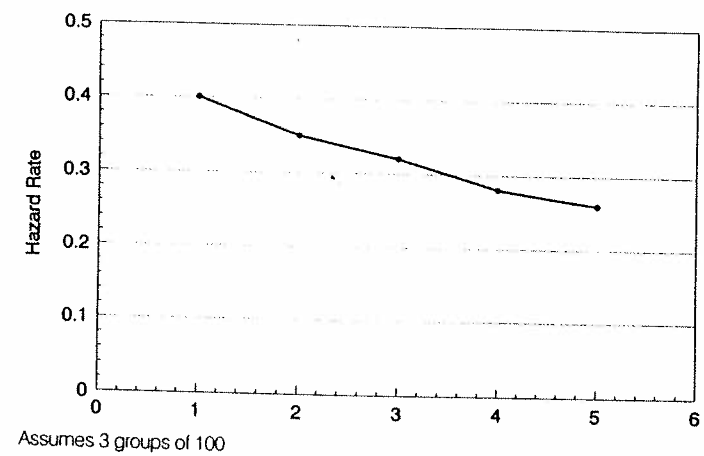


Fig. 8. Example of unobserved heterogeneity with three groups mixed together.

each of the groups. The hazard for Group 1 is 0.6, Group 2 is 0.4, and Group 3 is 0.2. If we analyze the subgroups separately, the correct hazard rates would appear as in Figure 7. However, if we analyzed the 300 persons together, the hazard rate would appear to decline over time, as shown by Figure 8. How does this happen? The individuals in Group 1, having a higher hazard, are more likely to exit the sample earlier. (And conversely, the individuals in Group 3 are likely to stick around longer.) Thus, as time progresses, the average hazard for the entire sample decreases. This decrease, however, is a completely spurious effect due to the inclusion of different subgroups. The form of the underlying hazard function is constant, despite appearances! (You can work out this example for yourself by taking the numbers as presented in Figure 7. At each time point 60% of the first group will drop out, 40% of the second group, and 20% of the third group. Recalculate the average overall hazard rate after each time point.)

Fortunately, it is impossible to get spurious estimates of *increasing* hazard rates (Allison, 1984). That is, a problem of unobserved heterogeneity always acts to decrease the hazard rate. This means that whenever there is evidence to suggest that the hazard rate is increasing over time (as in our example), you can proceed to modeling without worrying about sources of unexplained heterogeneity.

ADVANCED METHODS AND ISSUES

The goal of this primer has been to provide an introduction to the basic techniques of survival analysis that are broadly applicable to a wide variety of data and are readily available in several major statistical packages. However, recent advances in statistics, biostatistics, sociology, and related fields have extended these basic techniques in several ways. Specifically, these newer methods can be applied to data that are richer or more precisely defined than the type of data often used in survival analysis. The following discussion briefly defines each of these methods, and points the reader to more detailed treatments.

Time-Varying Covariates

The Cox semiparametric model can be extended to include time-varying covariates. Time-varying covariates are those variables that can change value during the study or observation period. Thus, in addition to such typical time-constant explanatory variables as gender, other variables such as losing a job can be included in the model. The interpretation of results is straightforward: When a time-varying explanatory variable has an effect on the hazard rate, it will either cause the hazard to increase or decrease. For example, the loss of a job might increase the hazard of substance abuse recidivism.

Inclusion of time-varying covariates increases data collection and management difficulties. Time-varying information must be collected contemporaneously with the event data, although in some cases it is possible to gather this data retrospectively. Also, data files must be prepared differently if they include time-varying explanatory variables. The major software packages have increasingly started to include support for time-varying predictors. (See Appendix for a review of the major software packages.) Yamaguchi (1991) and Blossfeld et al. (1989) have excellent discussions of the Cox model and time-varying explanatory variables, including sample data and computer programs. Other good treatments include Allison (1984), Tuma and Hannan (1984), Kalbfleisch and Prentice (1980), and Cox and Oakes (1984).

Competing Risks

In the above discussion, it is assumed that the event under consideration is of one type. That is, 'dropping out of GROW' is one specifically defined event, and all instances of it are exactly alike. However, there are many times when the event being modeled may be thought of as a collection of similar events. For example, an analysis of job changes may include job transfers, retirement, being fired, and being laid off as different types of one general event. There may be times when a more detailed analysis of these different types of job changes would be appropriate. These different types of job changes are examples of competing risks. An employee is at risk at any time for any of these job change events. They are competitive in the sense that although any of them theoretically may occur at any time, only one can occur at any particular instant. A person can be fired or be transferred, but not both.

Fortunately, although competing risks introduces some theoretical complexity, its technical management is relatively simple. The solution is to model each specific event type separately. Thus, one analysis would be performed for job transfers, one for being fired, and so on. The overall hazard model for any type of job change is simply the sum of the type-specific models (Allison, 1984; Blossfeld et al., 1989, Carrol & Mayer, 1986).

Repeating Events

Just as survival analysis can be extended to handle multiple event types, it also can be generalized to deal with repeatable events. Although dropping out of GROW is not repeatable (at least the way it was defined), many events of interest to social scientists are. For example, pregnancy, abortion, arrests, and psychiatric recidivism are all events that commonly happen more than once, even within relatively short time spans.

One approach to dealing with repeatable events is to include as an observation each period of time between events for any individual as a separate observation. Thus, the time from prison release to first arrest would be one observation, and the time from the first arrest to the second arrest (if any) would be another observation. The pooled observations yield a model that estimates effects for all of the arrest intervals at once. Allison (1984) provides a general discussion of the strengths and weaknesses of this approach, and Heckman and Borjas (1980) provide a detailed empirical example of modeling repeating events.

Sociology research has produced a set of models that are quite different from the hazard rate models discussed in this paper. These models include all of the above models as special cases, but instead of dealing with hazard rates, they focus on changes or transition rates between different states. A major advantage of these state change models is their ability to handle both repeating and multiple types of events simultaneously. Although no major statistics package can currently estimate these models directly, most parts of the models can be analyzed by typical survival analysis proportional hazards techniques. For more details see Coleman (1981), Tuma (1982), Tuma and Hannan (1984), and Tuma, Hannan, and Groeneveld (1979).

IMPLICATIONS FOR APPLIED PSYCHOLOGICAL RESEARCH

By now it should be clear that survival analysis is a powerful technique for understanding the timing and context of qualitative change. Survival analysis can handle different types of data (e.g., continuous, discrete, prospective, retrospective, and archival data) and can deal with censored observations without bias or loss of cases. Also, it can model events nonparametrically, semiparametrically, or fully parametrically. Finally, it can describe the survival and hazard rates for the entire sample or for subgroups, and it can characterize the effects of continuous or categorical covariates. In other words, using only a few pieces of easily collected data, a wealth of information about the form and process of an event can be discovered by survival analysis.

The examples described above illustrate the utility of these techniques. With just a few variables and fewer than six analyses we were able to discover quite a bit about the process of dropping out of GROW. The typical GROW member attended for 2.8 months, but almost 40% of GROW members drop out after only 1 month. For the entire sample, the probability of dropping out is greatest in the first 2 months and increases again after 18 months of membership. More specifically, for those members who attend at least three meetings, the risk of dropping out is approximately .10 each month for the first 18 months of membership. After this the hazard increases to around .15 to .20. GROW members who have never been married, are older, or are lower functioning are the most likely to stay involved with GROW longer.

I conclude this primer with a short list of suggestions and/or tips that might be useful for anyone interested in using survival analysis. These have been gleaned from other survival analysis texts, studies using survival analytic techniques, and from my own experiences with teaching and using these methods.

1. Remember that under most circumstances survival analysis is able to handle event data measured at discrete time intervals. This means that survival analysis can often be used with the traditional longitudinal designs familiar to most social scientists. (I am specifically excluding the pre–post design. Although it is longitudinal, it is barely so!) Even if continuous time models are applied to event data measured discretely, the time aggregation bias can be minimized with appropriate choice of the size of the time interval. In other words, using survival analytic tools should not require substantial changes in study designs or methods. (Other longitudinal methods such as time-series analysis often require such changes.)

2. The use of archival or retrospective data, too often shunned in our field, can dramatically increase the utility of survival analysis. Event timing information is often recorded accurately in external data bases, and therefore represents a treasure trove to the intrepid researcher. For example, schools and organizations often maintain detailed absentee records. These can be used along with other information obtained from students or employees to model attendance, absenteeism, or dropouts. Also, depending on the type of event and the methods used to elicit the information, respondent retrospective reports can be highly reliable.

3. Most of the major software packages can competently handle the majority of the analyses described in this paper. However, there are some important differences between the programs. For example, not all of the major statistics packages currently handle time-varying covariates. (See Appendix for a more detailed discussion of these issues.) Although this can be somewhat frustrating, the output from various packages is usually clear and easily interpretable.

4. Much of the most interesting substantive results from survival analysis is graphical in nature (the survival and hazard functions, the subgroup curves, etc.) This makes it much easier to quickly grasp the important aspects of the event timing. Pictures of survival and hazard patterns really are worth a thousand words, and investigators should exploit the advantages

of visual displays of this information (Tufte, 1983). However, the graphs and plots produced by many of the survival analysis statistics packages leave much to be desired. Indeed, it may be necessary to replot the results before an accurate interpretation can be made. In particular, the plots of subgroup analyses often only become clear when they are redone using a more sophisticated graphics package such as Harvard Graphics or Freelance Graphics (to name just a couple).

5. You should exploit the power of survival analysis to examine the form of the underlying hazard function. Although covariate analysis is straightforward using the Cox regression model, by its very nature it limits the focus to how one set of variables is related to the event. Our field has traditionally emphasized such covariate-based analyses. By expanding our theoretical and empirical focus to include analyses of relationships among actors (via network analysis or cluster analysis) or relationships of events over time (time-series, growth curves, and survival analysis) we can better understand phenomena as whole entities rather than dissected slices of variance (Rapkin & Luke, 1993).

This primer has intentionally ignored a number of complex and important statistical issues. There are numerous statistical treatments of survival analysis available to deal with these issues, if needed. My purpose here, instead, has been to show how survival analysis can be used in a simplistic (but not simple-minded) manner to explore the richness of event data.

Survival analysis is a powerful tool. If we start including it in our analytic toolbox we will be better able to understand events and the qualitative changes that they represent. In turn, we will be able to develop social and psychological theories that have time and change as integral aspects.

APPENDIX: COMPUTER PROGRAM CONSIDERATIONS

Investigators interested in survival analytic techniques will be happy to find that there are several statistical computer packages that offer powerful survival analysis procedures. This has not always been the case. Just a few years ago, Goldstein et al. (1989) reported that although the available PC-based survival analysis packages gave accurate computational results, these programs had a long way to go. In particular, they noted the extremely poor documentation, lack of program flexibility, lack of diagnostic features, and general difficulty of use.

Since then the quality of survival analysis software has improved noticeably. In this Appendix, I briefly discuss four major software packages that have comprehensive procedures to analyze event duration data. The four packages have been chosen (BMDP, SAS, SPSS, and SYSTAT) because they are the most widely available and most frequently used by social scientists. I caution the reader that although the information presented here is, to the best of my knowledge, accurate at the time of writing (early 1993), it may become outdated rather quickly.

Table A1 summarizes the survival analysis features found in the four statistics packages. At the top of the table the specific version and platform are listed for each program. Features often vary between platforms for the

Table A1. Comparison of Survival Analysis Computer Packages

	BMDP	SAS	SPSS	SYSTAT
Package version and platform	BMDP 1990; Mainframe version	SAS Version 6; Mainframe version	SPSS for Windows Release 5; IBM PC	SYSTAT 5.02; IBM PC
Procedure names	1L, 2L	LIFETEST, LIFEREG, PHREG	SURVIVAL, KM, COXREG	Survival 1.0
Breadth of features	Good	Excellent	Fair	Fair
Ease of use	Fair	Poor	Excellent	Good
Graphics	Fair	Fair	Good	Good
Documentation	Fair	Fair	Fair	Fair
Nonparametric analyses				
Survival estimates	Life Table; Product-Limit	Life Table; Product-Limit	Life Table; Product-Limit	Product-Limit
Aggregated data	Yes	No	Yes	Yes
Rank tests	3	2	3	3
Censoring & failure pattern analysis	Yes	Yes	No	Yes
Semiparametric analyses				
Cox proportional hazards	Yes	Yes	Yes	Yes
Stepwise models	Yes	Yes	Yes	Yes
Covariate vector models	Yes	Yes	Yes	Yes
Time-varying covariates	Yes	Yes	Yes	No
Parametric analyses				
Fully parametric models	Yes	Yes	No	Yes
Covariates	Yes	Yes	No	Yes
Stepwise models	Yes	No	No	Yes

same program. For example, currently in SPSS many of the advanced survival analysis procedures (e.g., Cox proportional hazards modeling) are only available in the Windows version, not in the mainframe or regular DOS versions. For each package the specific survival analysis procedure names are listed. In the case of SYSTAT, the survival analysis routines are part of a separate module that does not come with the basic SYSTAT package; it must be purchased separately. SPSS for Windows is currently provided in three separate packages; all of the survival procedures are found in the Advanced Statistics module.

General Features

After the procedure names, I rate the programs on breadth of features, ease of use, graphics, and documentation. These are simple subjective ratings, and are meant only to highlight the relative strengths and weaknesses of the packages. SAS clearly stands out with regard to the variety of things that can be done with the survival procedures. For example, SAS is the only procedure that allows the user to specify how event duration time ties should be handled by the program. However, perhaps because of the depth of the program, SAS is also the most difficult to use. SAS puts a lot of power into the hands of the analyst, but it is clearly not meant for the novice user. SPSS for Windows stands in sharp contrast. Its modern graphical interface, support for interactive data management and analysis, and speedy operation make it by far the easiest survival analysis package to use. (At the time of this writing SAS and SYSTAT both had just released Windows versions of their software. However, the SYSTAT Windows package did not include the survival analysis procedures, and the SAS package was not available to the author.)

Not surprisingly, graphics support is weaker in the mainframe packages. SPSS with its Windows environment makes it extremely easy to get quick graphics output of survival and hazard plots. SYSTAT, with its integrated Sygraph graphing facility, provides more powerful support, although it is somewhat harder to use.

Documentation remains a frustrating aspect of these programs. Although there are some strengths, the documentation is generally poorly structured and hard to understand. Although all provide programming examples, there could still be more coverage of tricky procedures, such as how to include time-varying covariates in the model. The documentation for SAS' LIFETEST and LIFEREG procedures is particularly dense and unhelpful. The PHREG documentation comes in the form of a technical report, and is much more thorough with many helpful examples complete

with data and program output (SAS Institute, 1990). It is still written, however, with the statistical professional in mind. The SPSS documentation has the most attractive presentation and is the easiest to read, but is only minimally helpful with how to do complicated analyses. The SYSTAT survival package comes with its own extensive 154-page manual. It is very readable and is the only set of documentation reviewed here that tries to educate the reader about survival analysis as well as describing the program. It includes an index and an adequate reference list. Despite its thoroughness the documentation seems unorganized at times; also, the examples could be more extensive. The BMDP documentation, on the other hand, is organized clearly but the presentation is too often cryptic.

Nonparametric Analyses

All four of the statistics packages competently produce nonparametric estimates of survival and hazard functions, and can perform rank tests for subgroups based on single categorical predictor variables. There are no major limitations in any of the implementations of these procedures. There are a few minor differences between the programs. All the programs except SAS can estimate survival and hazard functions for aggregated data. (Instead of having event duration times and censor information for each case, aggregated data are in the form of counts per time interval.) In addition to the overall test score for comparing survival curves of subgroups defined by a categorical predictor variable, SPSS can automatically compare each pair of categorical factor levels. This must be done manually in the other programs. SYSTAT lacks the ability to produce survival estimates based on the life table approach. Finally, all of the packages except SPSS provide simple plots of censoring and survival pattern times. These plots are useful for detecting if censoring is truly independent from survival times.

Semiparametric Analyses

Support for Cox proportional hazards modeling is particularly strong for all of these programs. In fact, except for SYSTAT's inability to handle time-varying covariates, these procedures look remarkably similar. Both SAS and SPSS offer a wide array of diagnostic plots, used to test for violations of the proportional hazards model. SPSS can provide these plots automatically, while in SAS they must be specifically programmed.

While support for time-varying covariates is one of the major strengths of the proportional hazards model, neither BMDP, SPSS, nor SAS make it easy to perform these analyses. For example, if you would like to include a time-varying predictor that is not a simple function of time, BMDP requires you to write a separate FORTRAN subroutine. SAS seems to have the best support for this type of analysis.

Parametric Analyses

All of the packages except for SPSS support a variety of parametric models in very similar ways. The documentation for SYSTAT provides a nice summary of how to interpret the various parametric coefficients such as shape and scale. None of the programs provide enough examples in their documentation for parametric models.

Summary

For the most part, any of the four statistical packages discussed here can be used to perform the types of survival analyses discussed in this primer. If you need straightforward estimates of survival and hazard functions, or simple tests of the effects of covariates on these estimates, any of these packages will suffice. Although the focus of this primer has been on continuous time survival models, it should be noted that the logit regression procedures in these statistics packages can also be used to perform discrete time hazard modeling.

There is still a lot of room for improvement. None of the packages at this time can automatically handle competing risks or repeating events. Support for time-varying covariates could be improved. It would be nice to see discrete time models integrated into the survival procedures, rather than having to turn to the general logit regression procedures. Finally, the documentation for all the packages could be improved.

REFERENCES

Adler, I., & Kandel, D. B. (1983). Risk periods for drug involvement in adolescence in France and in Israel: Application of survival analysis to cross-sectional data. *Social Forces, 62*, 375-397.

Allgulander, C., & Fisher, L. D. (1986). Survival analysis (or time to an event analysis), and the Cox regression model—methods for longitudinal psychiatric research. *Acta Psychiatrica Scandinavica, 74*, 529-535.

Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 61-98). San Francisco: Jossey-Bass.

Allison, P. D. (1984). *Event history analysis: Regression for longitudinal event data.* Newbury Park, CA: Sage.

Blossfeld, H., Hamerle, A., & Mayer, K. U. (1989). *Event history analysis.* Hillsdale, NJ: Erlbaum Associates.

Breslow, N. E. (1974). Covariance analysis of censored survival data. *Biometrics, 30*, 89-99.

Campbell, R. T., Mutran, E., & Parker, R. N. (1987). Longitudinal design and longitudinal analysis. *Research on Aging, 8*, 480-502.

Carroll, G. R., & Mayer, K. U. (1986). Job-shift patterns in the Federal Republic of Germany: The effects of social class, industrial sector and organizational size. *American Sociological Review, 51*, 323-341.

Coleman, J. S. (1981). *Longitudinal data analysis.* New York: Basic Books.

Collins, L. M. (1991). Measurement in longitudinal research. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 137-148). Washington, DC: American Psychological Association.

Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society. Series B, 34*, 187-202.

Cox, D. R., & Oakes, D. (1984). *Analysis of survival data.* New York: Chapman and Hall.

Cronbach, L. J., & Furby, L. (1970). How we should measure "Change"—or should we? *Psychological Bulletin, 74*, 68-80.

Dixon, W. J. (Ed.) (1990). *BMDP statistical software manual* (Vol. 2). Berkeley: University of California Press.

Dunn, G., & Master, D. (1982). Latency models: The statistical analysis of response times. *Psychological Medicine, 12*, 659-665.

Felmlee, D., Sprecher, S., & Bassin, E. (1990). The dissolution of intimate relationships: A hazard model. *Social Psychology Quarterly, 53*, 13-30.

Fergusson, D. M., Horwood, L. J., & Shannon, F. T. (1984). A proportional hazards model of family breakdown. *Journal of Marriage and the Family, 46*, 539-549.

Fichman, M. (1989). Attendance makes the heart grow fonder: A hazard rate approach to modeling anendance. *Journal of Applied Psychology, 74*, 325-335.

Fisher, D. G., & Anglin, M. D. (1987). Survival analysis in drug program evaluation. Part I. Overall program effectiveness. *International Journal of the Addictions, 22*, 115-134.

Gail, M. H., Santner, T. J., & Brown, C. C. (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics, 36*, 255-266.

Gehan, E. A. (1975). Statistical methods for survival time studies. In *Cancer therapy: Prognostic factors and criteria* (pp. 7-35). New York: Raven.

Goldstein, R., Anderson, J., Ash, A., Craig, B., Harrington, D., & Pagano, M. (1989). Survival analysis software on MS/PC-DOS computers. *Journal of Applied Econometrics, 4*, 393-414

Greenhouse, J. B., Kupfer, D. J., Frank, E., Jarrett, D. B., & Rejman, K. A. (1987). Analysis of time to stabilization in the treatment of depression: Biological and clinical correlates. *Journal of Affective Disorders, 13*, 259-266.

Greenhouse, J. B., Stangl, D., & Bromberg, J. (1989). An introduction to survival analysis: Statistical methods for analysis of clinical trial data. *Journal of Consulting and Clinical Psychology, 57*, 536-544.

Griffin, W. A., & Gardner, W. (1989). Analysis of behavioral durations in observational studies of social interaction. *Psychological Bulletin, 106*, 497-502.

Gruenewald, P. J., & West, B. R. (1989). Survival models of recidivism among juvenile delinquents. *Journal of Quantitative Criminology, 5*, 215-229.

Harrison, D. A., & Hulin, C. L. (1989). Investigations of absenteeism: Using event history models to study the absence-taking process. *Journal of Applied Psychology, 74*, 300-316.

Hays, W. L. (1963). *Statistics for psychologists.* New York: Holt, Rinehart and Winston.

Heckman, J. J., & Borjas, G. (1980). Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence. *Econometrica, 47*, 247-283.

Heijtan, D. F. (1989). Inference from grouped continuous data: A review. *Statistical Science, 4*, 164-183.

Hocking, R. R. (1983). Developments in linear regression methodology: 1959-82. *Technometrics, 25*, 219-230.

Hser, Y. I., Anglin, M. D., & Liu, Y. (1991). A survival analysis of gender and ethnic differences in responsiveness to methadone maintenance treatment. *The International Journal of the Addictions, 25*, 1295-1315.

Isager, T., Brinch, M., Kreiner, S., & Tolstrup, K. (1985). Death and relapse in anorexia nervosa: Survival analysis of 151 cases. *Journal of Psychiatric Research, 19*, 515-521.
Kalbfleisch, J. D., & Prentice, R. L. (1980). *The statistical analysis of failure time data.* New York: Wiley.
Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*, 457-481.
Katz, R. (1980). Time and work: Toward an integrative perspective. In B. Staw & L. Cummings (Eds.), *Research in organizational behavior* (Vol. 2, pp. 81-128). Greenwich, CT: JAI.
Kennedy, J. J. (1983). *Analyzing qualitative data: Introductory log-linear analysis for behavioral research.* New York: Praeger.
Kraft, P. (1991). Age at first experience of intercourse among Norwegian adolescents: A lifestyle perspective. *Social Science and Medicine, 33*, 207-213.
Lavori, P. W., Keller, M. B., & Klerman, G. L. (1984). Relapse in affective disorders: A reanalysis of the literature using life table methods. *Journal of Psychiatric Research, 18*, 13-25.
Lawless, J. F. (1982). *Statistical models and methods for lifetime data.* New York: Wiley.
Leon, A. C., Friedman, R. A., Sweeney, J. A., Brown, R. P., & Mann, J. J. (1990). Statistical issues in the identification of risk factors for suicidal behavior: The application of survival analysis. *Psychiatry Research, 31*, 99-108.
Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York:
Lovell, M. C. (1983). Data mining. *Review of Economics and Statistics, 65*, 1-12.
Luke, D. A., & Roberts, L. (1992, August). *A two-stage model of self-help group participation.* Paper presented at the annual meeting of the American Psychological Association, Washington, DC.
Luke, D. A., Roberts, L., & Rappaport, J. (in press). Individual, group context, and individual-group fit predictors of self-help group attendance. *Journal of Applied Behavioral Science.*
MacKenzie, D. L. (1991). The parole performance of offenders released from shock incarceration (boot camp prisons): A survival time analysis. *Journal of Quantitative Criminology, 7*, 213-236.
Mantel, N. (1966). Evaluation of survival data and two new rank.order statistics arising in its consideration. *Cancer Chemotherapy Reports, 50*, 163-170.
Matthews, S. M., Roper, M. T., Mosher, L. R., & Menn, A. Z. (1979). A non-neuroleptic treatment for schizophrenia: Analysis of the two-year postdischarge risk of relapse. *Schizophrenia Bulletin, 5*, 322-333.
McLeod, J. D., Kessler, R. C., & Landis, K. R. (1992). Speed of recovery from major depressive episodes in a community sample of married men and women. *Journal of Abnormal Psychology, 101*, 277-286.
Miller, R. G. (1981). *Survival analysis.* New York: Wiley.
Morita, J. G., Lee, T. W., & Mowday, R. T. (1989). Introducing survival analysis to organizational researchers: A selected application to turnover research. *Journal of Applied Psychology, 74*, 280-292.
Murnane, R. J., Singer, J. D., Willett, J. B. (1988). The career paths of teachers: Implications for teacher supply and methodological lessons for research. *Educational Researcher, 17*, 22-30.
Murnane, R. J., Singer, J. D., Willett, J. B., Kemple, J. J., & Olsen, R. J. (1991). *Who will teach? Policies that matter.* Cambridge, MA: Harvard University Press.
Petersen, T. (1991). Time-aggregation bias in continuous-time hazard-rate models. In P. V. Marsden (Ed.), *Sociological Methodology* (pp. 263-290). Washington, DC: American Sociological Association.
Peto., R. Pike, M. C., Armitage, P. Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K., Peto, J., & Smith, P. G. (1976). Design and analysis of randomized clinical trials requiring prolonged observation of each patient: Introduction and design. *British Journal of Cancer, 34*, 585-612.

Poikolainen, K. (1983). Survival methods in the evaluation of the outcome of alcoholism treatment. *British Journal of Addiction, 78*, 403-407.
Pollard, A. H., Yusuf, F., & Pollard, G. N. (1981). *Demographic techniques.* Sydney: Pergamon.
Quinn, R. E., & Cameron, K. S. (Eds.). (1988). *Paradox and transformation: Toward a theory of change in organization and management.* Cambridge, MA: Ballinger.
Rapkin, B. D., & Luke, D. (1993). Cluster analysis in community research: Epistemology and practice. *American Journal of Community Psychology, 21*(2).
Rappaport, J., Seidman, E., Toro, P. A., McFadden, L. S., Reischl, T. M., Roberts, L. J., Salem, D. A., Stein, C. H., & Zimmerman, M. A. (1985). Finishing the unfinished business: Collaborative research with a mutual help organization. *Social Policy, 16*, 12-24.
Roberts, L. J., Luke, D. A., Rappaport, J., Seidman, E., Toro, P., & Reischl, T. M. (1991). Charting uncharted terrain: A behavioral observation system for mutual help groups. *American Journal of Community Psychology, 19*, 715-737.
Russell, D. E. H., & Howell, N. (1983). The prevalence of rape in the United States revisited. *Signs: Journal of Women in Culture and Society, 8*, 688-695.
Ryan, J., & Dent, O. (1984). An introduction to survival analysis: Factors influencing the duration of breast feeding. *Australian and New Zealand Journal of Sociology, 20*, 183-196.
SAS Institute Inc. (1990). *SAS/STAT user's guide. Version 6* (Vol. 2, 4th ed.), Cary, NC: Author.
Setze, P. J., & Bond, G. R. (1985). Psychiatric recidivism in a psychosocial rehabilitation setting: A survival analysis. *Hospital and Community Psychiatry, 36*, 521-524.
Singer, J. D., & Willett, J. B. (1991). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin, 110*, 268-290.
Slonim-Nevo, V., & Clark, V. A. (1989). An illustration of survival analysis: Factors affecting contraceptive discontinuation among American teenagers. *Social Work Research and Abstracts, 25*, 7-14.
SPSS Inc. (1991). *SPSS statistical algorithms* (2nd ed.). Chicago: Author.
Sroufe, L. A., & Jacobvitz, D. (1989). Diverging pathways, developmental transformations, multiple etiologies and the problem of continuity in development. *Human Development, 32*, 196-203.
Steinberg, D., & Colla, P. (1988). *SURVIVAL: A supplementary module for SYSTAT.* Evanston, IL: Systat Inc.
Stokols, D. (1988). Transformational processes in people-environment relations. In J. E. McGrath (Ed.), *The social psychology of time* (pp. 233-252). Newbury Park, CA: Sage.
Tarone, R. E., & Ware, J. (1977). On distribution-free tests for equality of survival distributions. *Biometrika, 64*, 156-160.
Tufte, E. R. (1983). *The visual display of quantitative information.* Cheshire, CT: Graphics.
Tuma, N. B. (1982). Nonparametric and partially parametric approaches to event history analysis. In S. Leinhardt (Ed.), *Sociological Methodology 1982.* San Francisco: Jossey-Bass.
Tuma, N. B., & Hannan, M. T. (1984). *Social dynamics: Models and methods.* San Diego, CA: Academic Press.
Tuma, N. B., Hannan, M. T., & Groeneveld, L. P. (1979). Dynamic analysis of event histories. *American Journal of Sociology, 84*, 820-854.
Wicker, A. W., & King, J. C. (1988). Life cycles of behavior settings. In J. E. McGrath (Ed.), *The social psychology of time* (pp. 182-200). Newbury Park, CA: Sage.
Widaman, K. F. (1991). Qualitative transitions amid quantitative development: A challenge for measuring and representing change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 204-217). Washington, DC: American Psychological Association.
Wilkinson, L. (1979). Tests of significance in stepwise regression. *Psychological Bulletin, 86*, 168-174.
Wilkinson, L. (1990). *SYSTAT: The system for statistics.* Evanston, IL: Systat, Inc.
Willett, J. B. (1987). Questions and answers in the measurement of change. *Review of Research in Education, 49*, 587-602.
Willett, J. B., & Singer, J. D. (1991a). How long did it take? Using survival analysis in educational and psychological research. In L. M. Collins, & J. L. Horn, (Eds.), *Best*