

In practice, a value of time is obtained by calculating the number of days (or months, or years, etc.) between two calendar dates. Table 1.1 shows the admission date and the follow up date for the subjects in this sample from the WHAS study. Most statistical software packages have functions that allow the user to manipulate calendar dates in a manner similar to other numeric variables. They do this by creating a numeric value for each calendar date, which is defined as the number of days from some predetermined reference date. For example, the reference date used by most, if not all, packages is January 1, 1960. Subject 5 entered the study on February 9, 1995, which is 12,823 days after the reference date, and died May 29, 1998, which is 14,028 days after the reference date. The interval between these two dates is $14,028 - 12,823 = 1,205$ days. The number of days can be converted into the number of months by dividing by $30.4375 = (365.25 / 12)$. Thus, the survival time in months for subject 5 is $39.589 = (1,205 / 30.4375)$. It is common, when reporting results in tabular form, to round months to the nearest whole number, e.g., 40 months. The level of precision used in reporting and analyzing survival time should depend on the particular application.

Two mechanisms can lead to incomplete observation of time: censoring and truncation. A censored observation is one whose value is incomplete due to factors that are random for each subject. A truncated observation is incomplete due to a selection process inherent in the study design. The most commonly encountered form of a censored observation is one where observation begins at the defined time $t = 0$ and terminates before the outcome of interest is observed. Because the incomplete nature of the observation occurs in the right tail of the time axis, such observations are said to be *right censored*. For example, in the WHAS study, a subject could move out of town or still be alive at the last follow up. In a study where right censoring is the only type of censoring possible, observation on subjects may begin at the same time or at varying times. For example, in a test of computer life length, we may begin with all computers started at exactly the same time. In a randomized clinical trial or in an observational study, such as the WHAS study, patients may enter the study over several years. As we see in Table 1.1, subject 2 entered the study on January 14, 1995, while subject 50 entered on July 17, 1997. In this type of study, regardless of calendar time, each subject's time of enrollment is assumed to define the $t = 0$ point.

For obvious practical reasons, all studies have a point when observation ends on all subjects; therefore subjects entering at different times will have variable lengths of maximum follow-up time. In the WHAS study, the last follow up date is December 31, 2002. Subject 13 entered the study on May 21, 1995. Thus the longest this subject could have been followed is 7 years, 7 months, and 10 days. However, this subject was not followed for the maximum length of time because the subject died on March 18, 1996, yielding a survival time of 302 days. Incomplete observation of a survival time due to the end of the study or follow-up is considered a right censored observation because the process by which subjects

A typical pattern of entry into a follow-up study is shown in Figure 1.1. This is a hypothetical 2-year study in which patients are enrolled during the first year. We see that subject 1 entered the study on January 1, 1990, and died on March 1, 1991. Subject 2 entered the study on February 1, 1990, and was lost to follow-up on February 1, 1991. Subject 3 entered the study on June 1, 1990, and was still alive on December 31, 1991, the end of the study. Subject 4 entered the study on September 1, 1990, and died on April 1, 1991. Subjects 2 and 3 have survival times that are right-censored. These data are plotted on the analysis time scale, in months, in Figure 1.2. Note that each subject's time is plotted as if he or she were enrolled at exactly the same calendar time and were followed until his or her respective end point. The two figures illustrate the difference between collecting data in calendar time and then converting it to analysis time.

In some studies, there may be a clear definition of the beginning time point; but subjects may not come under actual observation until after this point has passed. For example, in modeling age at menarche, suppose we define the zero value of time as 8 years. Suppose a subject enters the study at age 10, still not having experienced menarche. We know that this subject could have experienced menarche after age 8 but, due to the study design, was not enrolled in the study until age 10. This subject would not enter the analysis until time 10. This type of incomplete observation of time is called *left truncation* or *delayed entry*. Another example would be to study survival time in the WHAS among those discharged from the hospital alive. Here subjects stay in the hospital for varying lengths of time but we do not begin to study them until they "leave the front door."

