

Statistics for Nuclear Medicine

Part 5: Survivorship Studies

Peter C. O'Brien, Marc A. Shampo, and James S. Robertson

Mayo Clinic and Mayo Foundation, Rochester, Minnesota

J Nucl Med 24: 444-446, 1983

Definition of study group. In every study of survivorship—as in virtually all medical research on human subjects—the first requirement is to describe the group studied. The reader of the report must be told the nature of the group so he can judge whether his patient or group is like it. The description should include:

1. The source of subjects and the period in which they entered the study, with notice of any considerable selection bias (practice in a general hospital or a specialty clinic, and so forth).
2. The medical problem of interest: what it was and how its presence was determined. In some studies it is desirable to distinguish subtypes of the problem or degrees of severity.
3. The treatment, if any.
4. All exclusions of subjects from the study and the reasons for them.
5. Characteristics of the study group: their age and sex distributions; if pertinent, their area of residence, occupations, economic status, race, and so on.
6. Complicating features (associated diseases, et cetera) if it seems they may affect survival.

Data collection and accounting. *Completeness of follow-up.* The problem in follow-up is the practical difficulty of making it complete enough. Much effort and many stratagems may be justified, because a case “lost to follow-up” cannot be ignored. Even if entirely excluded from the analysis, it must be mentioned in the report and remembered in judgment, because cases lost may not have had the same outcome as the cases traced. No amount of sophisticated mathematical manipulation can overcome failure of follow-up in a sizable number of instances.

Initial event. In survivorship studies, each case must have an initial event from whose date the period of ob-

servaion is counted. This may be birth, for congenital disease; but usually it is diagnosis, surgery, or beginning of other treatment. Although the time of onset of the disease might be very meaningful, dating of onsets is often difficult. Surgical and hospital deaths may be excluded (if exclusion is desired) by beginning at a time such as “30 days after operation.”

Accounting of follow-up period. Since the initial event does not ordinarily occur simultaneously in all cases, the lengths of follow-up are not equal at any given date. Survivorship analysis, however, is based on an equal follow-up interval, which is attained at different times, case by case. A subject becomes eligible for inclusion in analysis of survival for a given period when that much time has passed since the initial event in his case. Thus a patient whose cancer was resected 3 yr ago is eligible for inclusion in analysis of 3-yr survival, despite having died of recurrence 2 yr after the resection. In two more years he will become eligible for 5-yr analysis; but he is not eligible for it now, even though we know now what his status will be then. To advance a 3-yr nonsurvivor to the 5-yr calculation would unbalance it, because we do not know how to advance (as alive or dead?) the 3-yr subjects presently surviving, who must be considered with him.

Data collected. The minimum amount of information on each subject for routine statistical analysis is listed in Table 1.

Analysis of data. *Direct (ad hoc) analysis.* Direct determination of a survival rate is done with this formula:

$$\frac{\text{Subjects who survived through the period of observation}}{\text{Subjects who survived that long plus those who were eligible but died}}$$

Single-period. Some years ago, it was usual to analyze survival data for the 5-yr rate alone. For example, if

Received Sept. 30, 1982; revision accepted Sept. 30, 1982.

For reprints contact: Dr. O'Brien, section of Medical Research Statistics, Mayo Clinic and Mayo Foundation, 200 First St., SW, Rochester, MN 55905.

TABLE 1. MINIMAL INFORMATION TO BE RECORDED FOR STUDY OF SURVIVORSHIP*

1. Sex
2. Date of birth (to give age at initial event)
3. Date of initial event
4. Date of latest follow-up (of death, if dead)
5. Status at latest follow-up (dead or alive)
6. Cause of death (if available)

* From O'Brien PC, Shampo MA: Statistics for clinicians:

11. Survivorship studies. *Mayo Clin Proc* 56:709-711, 1981. By permission.

gastrectomy had been performed on 84 patients five or more years previously, and at 5 yr after operation (case by case) there were 42 surviving, the 5-yr survivorship was $42/84 = 50\%$.

Single-period analysis, however, has two major inadequacies. First, a single-period rate does not reveal survivorship at any time preceding or following the end of the period chosen. Second, a single-period analysis (unless the period is brief) excludes a great deal of data an investigator is likely to have accumulated from more recent cases.

Serial determinations. It is possible, of course, to perform direct-method calculations on periods expanding from the initial event (1-yr, 2-yr, 3-yr, for instance—not first-year, second-year, third-year), each time, using all the cases eligible for the period being considered then. These serial determinations should reveal any trend within the maximal period analyzed.

The resultant series of rates, however, may not be very accurate. Indeed, if there has been less mortality among early cases than recent ones, this method may produce survival rates that rise with the length of follow-up. Some degree of such distortion may be present without being obvious, and each determination still excludes data from the computations. Therefore, this method is often not a good choice. For a more detailed nontechnical discussion, see Berkson and Gage (1).

Actuarial (life-table) analysis. Typically more accurate than the direct method is the actuarial method. This is based on the question, applied to each day of observation n ($n = 1, 2, \dots$), "For subjects who survived n days, what is the probability (p_n) of surviving one more day?" (To estimate this probability, we divide the number of subjects who actually survived $n + 1$ days by this number plus the number who died on the $n + 1$ st day.) The probability of surviving from day 1 through day n is then estimated by the product of the probabilities of surviving each day ($p_1 \cdot p_2 \dots p_n$). Although the computations for this method may appear cumbersome in computing a 5-yr survival rate, they are greatly simplified by the fact that, except for days on which deaths

occurred, $p_n = 1$. (Also for very large data sets, the computations are performed by computer.)

The major advantage of the actuarial method is that it utilizes all the available data: every subject is counted for whatever time he has been followed, no matter how brief. This makes the estimated survival rates more reliable. Second, the rates for successive intervals are combined in a way that excludes distortion. A curve that makes survival appear to increase as time passes is not possible.

Deaths due to unrelated causes. Thus far we have described determination of the gross death rate among a study group. If any of the deaths were due to causes other than the risk factor under study, however, and if the investigator is sure of his knowledge in every case, he must decide whether to determine and report the cause-specific death rate. This is accomplished by treating as deaths only those instances caused by the risk factor. Unrelated deaths are treated as lost to follow-up at time of death. Usually the particular study dictates the greater interest, and sometimes both rates are of interest.

Presentation of results. Generally the most effective method for describing the survival experience of a group of patients is to graph survival rates against time as shown in Figure 1.

To provide perspective on the outcome of an analysis, a comparison with normal survivorship may be shown. The appropriate norm is experience in a segment of the general population, adjusted (from published tables) to match the study group with respect to age, sex, and perhaps other features that seem pertinent. These rates will indicate the survivorship that would have been expected in the study group if it were representative of the general population. Moreover, expected 5-yr or 10-yr rates might be presented in the text.

Comment. The principal concern of this paper is to point out the need to take varying lengths of follow-up

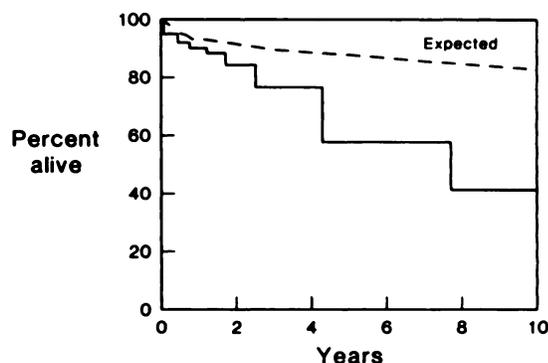


FIG. 1. Survivorship (actuarial analysis): as observed in study group (—) and as derived from population segment similar in regard to age, sex, and date of birth (---). (From O'Brien PC, Shampo MA. Statistics for clinicians. 11. Survivorship studies. *Mayo Clin Proc* 56: 709-711, 1981. By permission.)

into account in studying survivorship. We hope that has received sufficient emphasis above.

Two other ideas remain for presentation here.

1. The methods for analyzing survival data have been developed more recently than the other statistical methods we have presented, and still newer techniques are being proposed continually. Procedures are available for testing the differences between two or more survival curves, for testing the association between survival and a continuous variable (such as ventricular ejection fraction), and for performing such tests after adjustments for other relevant factors.

2. Interpretation of results is often difficult, however, because survivorship studies generally are *observational* rather than experimental, and questions arise regarding what has caused the differences that are found.

To illustrate, suppose that two different surgical techniques were used to treat patients having the same disease and that 10-yr follow-up was obtained on all patients treated with each method. It would be tempting to attribute any difference in survivorship to the difference in surgical techniques. Such a conclusion might not be valid, however, since the disparity could be a result of other factors. For example, the two groups of patients may have been dissimilar with respect to factors that influence the choice of surgical technique (possibly severity of the illness or age of the patient). Unfortunately, sophisticated statistical algorithms are of only limited

usefulness in attempts to distinguish effects due to the factor of interest (surgery) from effects due to other causes.

In order to establish the relative merits of the two surgical techniques, it would be best to design an experiment specifically with this purpose in mind. Ideally, patients would be assigned randomly to either method, enabling a statistician to make a valid probability statement in comparing the two procedures.

Notice that this was the approach in the *experimental* studies described previously in this series. For example, the experimental study (Comparing two samples) was designed carefully, in advance of data collection, so that a direct comparison could be made of the change in free thyroxine by each of the two regimens used. When a difference between regimens is observed in a properly designed experimental study, we can make a valid probability statement regarding the hypothesis that it was caused entirely by other factors instead. Thus, although an observational study is often considerably more convenient and less expensive than a carefully designed experiment, one must also consider the quality and interpretability of the results ultimately to be obtained.

REFERENCE

1. BERKSON J, GAGE RP: Calculation of survival rates for cancer. *Proc Staff Meet Mayo Clin* 25:270-286, 1950