Counting Recurrent Events in Cancer Research

Robert J. Glynn, Julie E. Buring

In clinical research, outcomes of interest often recur in the same patient. Examples include asthma attacks, urinary-tract infections, migraines, injuries, seizures in epileptics, and admissions to a hospital. Examples in cancer research include recurrent mammary tumors (1), skin cancers (2), and fractures in patients with cancer metastatic to bone (3). However, appropriate statistical methods to compare the occurrence of recurrent events between treatment groups are not widely understood or are they broadly disseminated to the research community. In this issue of the Journal, Cook and Major (4) illustrate the errors that can occur with inappropriate analytic strategies and propose some principled alternatives.

In the setting of recurrent events, research studies commonly count only the first occurrence of an outcome in a subject. For example, in the first publication of data from the pamidronate trial in women with breast cancer (5), used by Cook and Major to illustrate their methods, the primary analysis compared the time until a first important skeletal complication between the active and placebo groups. This analytic strategy is straightforward and unbiased, and it avoids the methodologic complications that can occur if a first event affects either risk of a subsequent event or compliance with ongoing randomized treatment in the trial (6). For example, occurrence of a first myocardial infarction markedly increases the risk of a subsequent event and can lead patients in the placebo arm of a randomized trial to initiate active therapy (7).

However, consideration of only first events may lead to an inaccurate evaluation of the efficacy of a treatment. In particular, it can substantially underestimate potential benefits in terms of events prevented by a treatment. In a variety of settings, ranging from studies of health care utilization (8) to recurrent falls in high-risk persons (9), an appropriate analysis of recurrent events can yield clinical insights and an understanding of public health impact that would be missed in an analysis of only first events. Furthermore, studies that discard relevant information on subsequent outcomes can require far greater sample sizes than studies that appropriately use all available outcomes on study subjects.

The major challenge to an appropriate analysis, including multiple events per person, arises because repeated events in the same person are almost never independent of one another. As noted by Cook and Major, the usual analysis of the “events per person-year” method assumes that all events occur randomly within each treatment group. More than 80 years ago, Green-
wood and Yule (10) presented a series of examples illustrating the principle that, across many settings in medical research, some people are more prone to recurrences of disease than others, and the assumption of independence of recurrent events is generally untenable. Each person may have an individual propensity for recurrent events and it may not be possible to measure variables that can distinguish between the differing rates for recurrences in individuals. The graphic display of individual fracture rates by Cook and Major demonstrates that, within the group of placebo-treated women, there is great variability, with some women remaining free of fractures throughout follow-up and others having frequent events. The poor fit of the simple model assuming a Poisson distribution indicates that this variability is far greater than what would be expected by chance. It was already clear from the work of Greenwood and Yule that an analysis of such data that ignores the potential for subjects to have individual propensities for recurrences would substantially underestimate the true variability in rates of recurrent events and lead to confidence intervals that are too narrow and $P$ values that are too small.

One may wonder why the principles demonstrated by Greenwood and Yule so long ago, and reiterated in many other examples of recurrent events since then, have often been ignored. Perhaps the problem arises because most basic textbooks in statistics pay little attention to examples of recurrent events, and dissemination of appropriate statistical software for analysis of recurrent events has been slow. The report by Cook and Major should help raise awareness of the analytic issues in the setting of recurrent events and suggest valid options.

Two other aspects of the data presented by Cook and Major further complicate analysis: Women in the trial were followed for varying lengths of time, and women with shorter follow-up times, who were nearer to death at randomization, had higher fracture rates. Several authors have presented methods for the analysis of recurrent events when persons are followed for varying intervals. One straightforward approach is to use the observed number of events in the population divided by the person-years of follow-up as a measure of the event rate but to use the variability in individual rates of occurrence to better estimate the variance of this rate (11,12). Because confounding by other variables is often a concern, regression models have also been developed for recurrent events (1,2,13–17) and applied to various settings, including clinical trials of seizures in epileptics (14) and variability in rates of malpractice claims (17). Recent books have summarized the available approaches (18,19). The uniqueness of the approach of Cook and Major is its consideration of follow-up time as a potential confounding variable. While the control of variables measured after randomization and potentially related to response must be carefully considered in terms of the potential for bias (20), the model of Cook and Major provides an improved explanation of the variability of event rates over time and between treatment groups.

Taken together with the abundance of evidence from other areas of medical research, the report by Cook and Major leads us to conclude that the assumption of randomly recurring events should not be used, unless accompanied by strong evidence that it fits the data well. Most often in the setting of recurrent events, some persons have greater propensities for recurrences than others, and comparisons of rates of recurrences between groups must take this into account. Several appropriate statistical approaches are available, although identification of the model that best fits the data can be challenging. However, the appropriate analysis can often answer questions of greater clinical relevance than a comparison of rates of first events.

REFERENCES


Note

Supported by Public Health Service grant EY08103 from the National Eye Institute, National Institutes of Health, Department of Health and Human Services.