MINIREVIEW

Hazard Ratio in Clinical Trials

Spotswood L. Spruance,[1]* Julia E. Reid,[1]† Michael Grace,[3] and Matthew Samore[1,2]

*Divisions of Infectious Diseases[1] and Clinical Epidemiology,[2] Department of Internal Medicine, University of Utah
School of Medicine, Salt Lake City, Utah, and Departments of Dentistry and Radiology, Faculty of
Medicine and Dentistry, University of Alberta, Edmonton, Canada[3]*

Time-to-event curves analyzed by Cox proportional hazards regression are commonly used to describe the outcome of drug studies. This methodology has the advantage of using all available information, including patients who fail to complete the trial, such as in cancer chemotherapy or human immunodeficiency virus antiviral treatment studies. The goal of treatment in such studies may be to prevent the development of a complication, for example, *Pneumocystis carinii* pneumonia, and to describe the likelihood of this complication's developing in the treatment group compared to the control group. The hazard ratio describes the relative risk of the complication based on comparison of event rates.

Hazard ratios have also been used to describe the outcome of therapeutic trials where the question is to what extent treatment can shorten the duration of the illness. However, the hazard ratio, a type of relative risk, does not always accurately portray the degree of abbreviation of the illness that occurred. In these circumstances, time-based parameters available from the time-to-event curve, such as the ratio of the median times of the placebo and drug groups, should be used to describe the magnitude of the benefit to the patient. The difference between hazard-based and time-based measures is analogous to the odds of winning a race and the margin of victory. The hazard ratio is the odds of a patient's healing faster under treatment but does not convey any information about how much faster this event may occur.

We have observed that there is substantial confusion among clinicians and clinical investigators about the difference between the hazard ratio and the median ratio. This report presents examples of this confusion, the exact distinction between the two statistics, and proper use of the hazard ratio and median ratio when interpreting the results of clinical trials to patients.

COX PROPORTIONAL HAZARDS MODEL

Time-to-event analysis. Clinical trials commonly record the length of time from study entry to a disease endpoint for a treatment and a control group. These data are commonly depicted with a Kaplan-Meier curve (Fig. 1), from which the median (time at which 50% of cases are resolved) and the mean (average resolution time) can be derived. The groups are compared by a time-to-event analysis (survival analysis) (1, 7). Time-to-event analysis provides a method to include patients who fail to complete the trial or do not reach the study endpoint (censored data) by making comparisons between the number of survivors in each group at multiple points in time. The alternative approach, excluding patients who are lost to follow-up, may introduce considerable bias because the data that these patients generate prior to their exit are important to the power and validity of the study. Time-to-event analysis can also incorporate information about subjects that may change over time (time-dependent covariates). Clinically important differences in the effect of treatment may be obscured if the proportions of survivors or recovered individuals in the treatment group are simply compared to that of the control group at a single point in time, such as at the conclusion of the trial (7). Time-to-event analysis is therefore a potentially more powerful and informative method of analysis.

Cox proportional hazards model and hazard ratio. There are several methods available to analyze time-to-event curves, such as Cox proportional hazards, log-rank, and Wilcoxon two-sample test, for example. The Cox proportional hazards model has been the most widely used procedure over many years of experience in medical research because of its applicability to a wide variety of types of clinical studies (2, 3).

The Cox model, a regression method for survival data, provides an estimate of the hazard ratio and its confidence interval. The hazard ratio is an estimate of the ratio of the hazard rate in the treated versus the control group. The hazard rate is the probability that if the event in question has not already occurred, it will occur in the next time interval, divided by the length of that interval. The time interval is made very short, so that in effect the hazard rate represents an instantaneous rate. An assumption of proportional hazards regression is that the hazard ratio is constant over time. Thus, in a clinical trial where disease resolution is the endpoint, the hazard ratio indicates the relative likelihood of disease resolution in treated versus control subjects at any given point in time.

USE OF THE COX MODEL IN TRIALS OF DISEASE DURATION

Interpretation of hazard ratio with time-based endpoints. The hazard ratio derived from the Cox model does not translate directly into information about the duration of time until

---

* Corresponding author. Mailing address: Division of Infectious Diseases, Rm. 4B322, Department of Internal Medicine, University of Utah School of Medicine, 30 North 1900 East, Salt Lake City, UT 84132-2405. Phone: (801) 581-8811. Fax: (801) 585-6422. E-mail: woody.spruance@hsc.utah.edu.
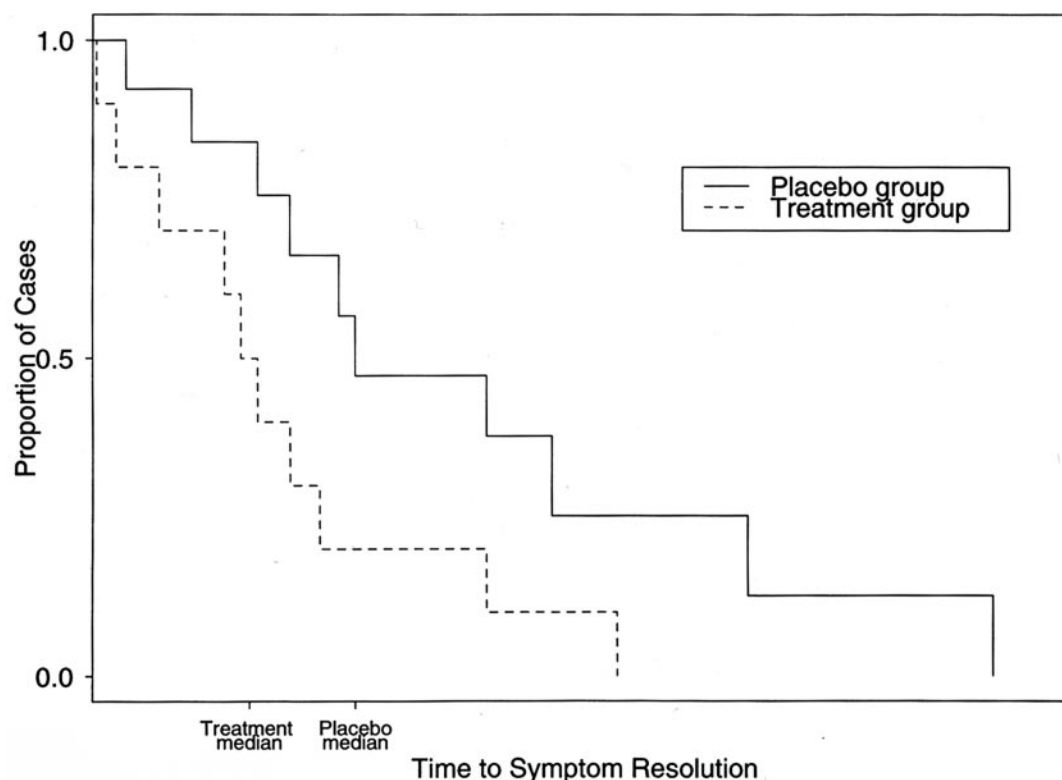† Present address: Myriad Genetics, Inc., Salt Lake City, UT 84108.

FIG. 1. Kaplan-Meier curve. Time to symptom resolution is compared between treatment and placebo groups. HR, hazard rate ratio = treatment hazard rate/placebo hazard rate. The hazard ratio is constant under the Cox proportional hazard model. The P value is used to reject the null hypothesis that HR = 1, i.e., treatment is not beneficial. Median, time at which half the cases are resolved and half are not resolved. MR, median ratio = placebo median time/treatment median time. Mean, average resolution time (based on the area under the Kaplan-Meier curve).

events. If the hazard ratio indicates a beneficial treatment effect, this implies that the time to the endpoint was reduced by treatment. However, as found in the clinical trial literature, the magnitude of the hazard ratio may be greater or less than the treatment benefit apparent from the median endpoint times. To compare these two statistics, the median endpoint time ratio can be calculated by dividing the control group median value by the treatment group median value. If there were a direct relationship between hazard ratio and median ratio, in a successful trial where the hazard ratio was 2, the median ratio would also be 2, indicating that the time to the endpoint in the treated group was half that of the control group.

Examples of such an outcome can be found. In Fig. 2B, a herpes zoster treatment study examining the effect of antiviral therapy on the duration of pain, the hazard ratio was 1.9 while the median ratio for the placebo and the 750-mg famciclovir groups was 119/61 days (2.0) (9). However, in Fig. 2A, a genital herpes treatment study, the hazard ratio was 1.9 when the ratio of median episode lengths for the placebo and the 500-mg valaciclovir groups was 5.9/4.0 days (1.5) (8). The reverse may also be seen. In a study of the effect of antiviral treatment on the duration of herpes zoster pain (Fig. 2C), the hazard ratio was 1.5 when the median ratio for the placebo and acyclovir groups was 62/20 days (3.1) (6, 11).

We have observed substantial confusion among clinicians as to the meaning of hazard ratio. For many clinicians, hazard ratio is a relative speed. Words found in the literature that describe the effect of treatments on the resolution of viral diseases when the hazard ratio was significantly greater than one have included "accelerated time [hazard ratio shown]," "resolved [hazard ratio shown] times faster," "hazard ratios indicate a 1.3 to 1.5-fold faster time," "more than twice as fast [hazard ratio = 2.13]," and "healing time was 15% shorter [hazard ratio = 0.85]" (9–12).

Accordingly, a hazard ratio of 2.0 may be misinterpreted as showing that patients in the treated group healed twice as fast as those in the control group. Twice as fast could mean to the clinician that the median healing time was cut in half by the treatment; that twice as many patients were likely to have healed *on* a particular day; that twice as many patients were likely to have healed *by* a particular day; or that the treatment group was likely to have healed twice as rapidly as the control group. None of these interpretations is correct. Clinicians may confuse velocity, the amount of distance traveled per unit of time, and the hazard, the rate of events per person-time. While velocity can be measured in a single object based on its distance traveled during a period of time, the hazard rates can only be inferred in a probabilistic sense from the occurrence of events in a population of at-risk individuals during a follow-up time interval.

A hazard ratio of 2 does means that treatment will cause the patient to heal faster, but in a very specific sense. In the context of hazard ratio, "fast" means that a treated patient who has not yet healed by a certain time has twice the chance of being healed at the next point in time compared to someone in the control group. Since this definition is markedly different from
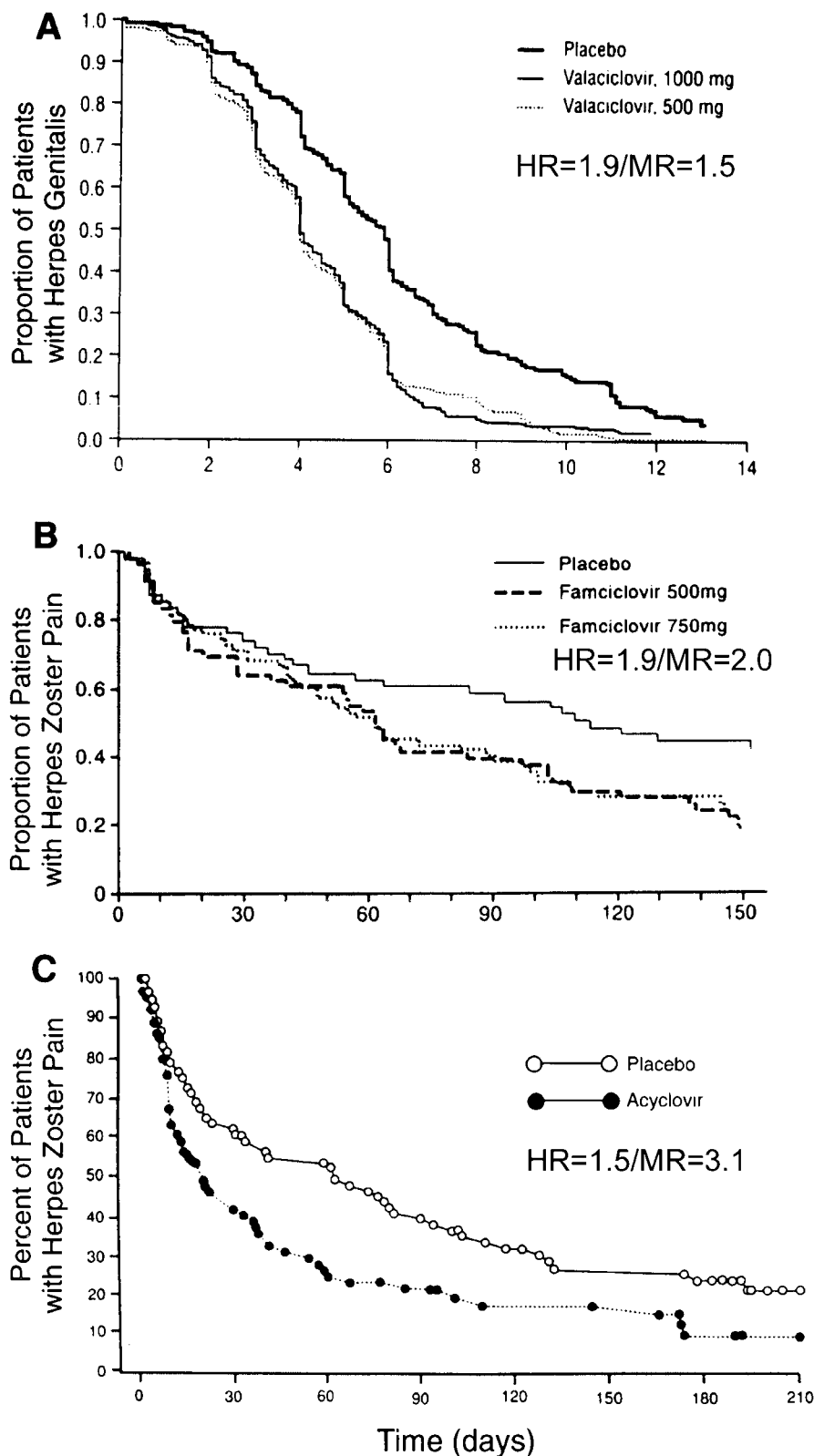
FIG. 2. Examples of the variable relationship between the hazard ratio and median ratio with clinical trials of antiviral agents as examples. While more than one drug dose was tested in two of the examples, outcomes were similar, and the statistics shown could apply to either curve. (A) Effect of valaciclovir on length of recurrent herpes genitalis episodes (8); hazard ratio = 1.9 and median ratio = 1.5. (B) Effect of famciclovir on time to resolution of postherpetic neuralgia (9); hazard ratio = 1.9 and median ratio = 2.0. (C) Effect of acyclovir on time to resolution of postherpetic neuralgia (6, 11); hazard ratio = 1.5 and median ratio = 3.1. The graphs were adapted from the indicated references.
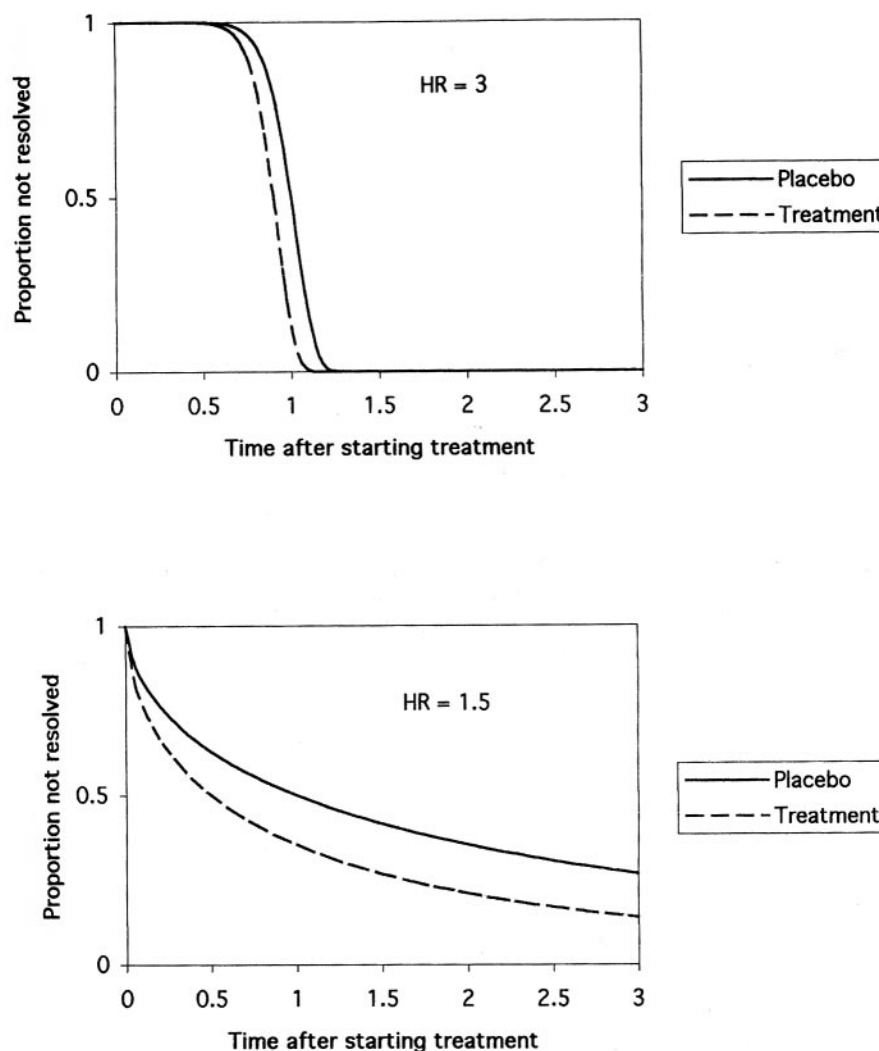
FIG. 3. The hazard ratio can be misleading if used to assess the amount of treatment benefit. Benefits of therapy depend not only on the hazard ratio but also on the shape of the underlying probability distribution, which is disease related. The survival functions in this figure are Weibull distributions that are standardized so the median survival time of the placebo group is one time unit. Mean survival times for the placebo groups are 1.0 (top figure) and 2.9 (bottom figure) time units. (Top) A relatively large hazard ratio can yield small treatment effects. These sinusoidal survival curves have a hazard ratio of 3, but both median and mean survival times are reduced only 10% by treatment. Although treatment increases the hazard rate, the effect is concentrated over such a limited time period that the overall benefit is minimal. (Bottom) A relatively small hazard ratio, in contrast, can yield large treatment effects. These concave survival curves have a hazard ratio of only 1.5, but both median and mean survival times are reduced 50% by treatment.

any intuitive meaning of the word fast described above, the hazard ratio can be misinterpreted. As illustrated in Fig. 2, reliance solely on the hazard ratio can lead to major errors in assessing treatment benefit in clinical trials. This is because the benefits of therapy depend not only on the hazard ratio but also on the shape of the underlying probability distribution, which is disease related (Fig. 3). Another way to express the distinction between hazard ratio and time-based measures of association is to consider the hazard as corresponding to effects on the vertical axis of a survival curve and time as corresponding to effects on the horizontal axis.

Hazard ratio, odds, and probability of healing. There is an alternative interpretation of the hazard ratio that may be intuitively easier to understand. The hazard ratio is equivalent to the odds that an individual in the group with the higher hazard reaches the endpoint first. Thus, in a clinical trial examining time to disease resolution, it represents the odds that a treated patient will resolve symptoms before a control patient. Stated another way, for any randomly selected pair of patients, one from the treatment group and one from the control group, the hazard ratio is the odds that the time to healing is less in the patient from the treatment group than in the patient from the control group. With the following equation, the probability of healing first can easily be derived from the odds of healing first, which is the probability of healing first divided by the probability of not healing first: hazard ratio (HR) = odds = $P/(1 - P)$; $P = HR/(1 + HR)$. A hazard ratio of 2 therefore corresponds to a 67% chance of the treated patient's healing first, and a hazard ratio of 3 corresponds to a 75% chance of healing first.

When the hazard ratio is thought of as the odds that a patient will heal faster with treatment, a unitless term not directly reflective of the fundamental time units of the study, it also becomes more evident that the hazard ratio cannot convey information about how much faster this event may occur. The difference between hazard-based and time-based measures is analogous to the odds of winning a race and the margin of victory.

QUESTIONS ASKED BY PATIENTS

A patient has recurrent herpes genitalis and is interested to know how episodic treatment with a new antiviral drug will improve the signs and symptoms of his recurrences. The physician has a copy of the results of a pertinent clinical trial (8). Cox proportional hazards regression was used to analyze the healing time, and a hazard ratio of 1.9 (95% confidence interval, 1.6 to 2.3) was reported for the lower of two drug doses tested. The median healing time was reduced by approximately 2 days, or 33%. The Kaplan-Meier curve is shown in Fig. 2A. The questions and the recommended responses by the physician are followed by a brief discussion. A fundamental challenge for the physician is to understand that the patient is asking about him- or herself, while the data by which the physician answers his or her questions derives from a population.

Q: Doctor, does this drug really work?

A: Yes, a clinical study has shown that the new drug promotes healing.

The hazard ratio is a clinical trial statistic that allows the physician to say with confidence that healing is faster with the new drug. The hazard ratio must be >1 and the lower limit of the 95% confidence interval of the hazard ratio must be >1, which was the case in this example.

Q: Doctor, what are the chances I will do better on this new drug compared to no treatment?

A: The odds are roughly 2:1 (the probability is 66%) that you will have an episode of shorter duration than someone who did not take the drug.

The odds are equal to the hazard ratio, which is 1.9 in the present case. The probability of healing sooner can be derived from the hazard ratio by the following formula: HR = odds = $P/(1 - P)$; $P = HR/(1 + HR)$. And so, in this example, $P = 1.9/2.9 = 0.67$.

It is unfortunately not absolutely certain that you will heal faster on this drug. There are two factors that influence the lesion healing time: the effect of the drug and natural variation in episode severity, so that some lesions will be very brief among untreated patients.

Q: Doctor, when will I heal if I use the new drug?

A: The study showed that about half the people who used the new drug healed within 4 days, and 95% healed within 8 days. Your experience will vary, like those of the people in the study, because of the natural variation in severity characteristic of this illness.

The question of timing can be expressed in several different ways. How likely is it that I will heal in a certain number of days if I use this treatment? By what day will I have a certain likelihood of healing? Answers to these questions can be determined from the Kaplan-Meier curve (see Fig. 2A). Healing times cannot be deduced from the hazard ratio unless the data are fit to an underlying parametric survival distribution (4, 5).

Q: Doctor, how much good will this drug do for me?

A: The new drug reduced the healing time in the study compared to the placebo group by about a third. Drug-treated patients had a median healing time of 4 days, compared to 5.9 days for the placebo recipients. If you use the new drug properly over a period of time, you can expect approximately this amount of benefit in comparison to what might have happened if you had let your lesions go untreated.

Reduction in the healing time can be estimated from the median healing times of the treated and placebo groups as shown in the survival curve. Here the patient is thinking of himself as two persons, one who takes the drug for recurrent herpes episodes over a period of time and one who does not take the drug. If taken properly, the drug should provide to an individual, compared to the hypothetical situation where one does not take the medication, approximately the same benefit calculated in the study from the difference between the healing times of the two treatment groups. This is the closest we can come to personalizing clinical trial results when discussing prospective treatment with a patient. While there may be pathophysiologic reasons for believing a patient will receive a certain degree of benefit every time the medication is taken, clinical data only allow comparisons between groups of individuals. As pharmaceutical company commercials on television warn, "individual results may vary."

CONCLUSIONS AND RECOMMENDATIONS

The Cox proportional hazards model is an appealing analytic method because it is both powerful and flexible. The hazard ratio, which is derived from this model, provides a statistical test of treatment efficacy and an estimate of relative risk of events of interest to clinicians. Examples of situations where the risk of an event is the question include the development of *Pneumocystis carinii* pneumonia in human immunodeficiency virus-infected patients, coronary reinfarction following stent placement, breast cancer in patients on estrogen supplements, and cardiovascular morbidity in patients taking aspirin.

However, the hazard ratio must be interpreted judiciously in clinical trials where the duration of events or the disease is the primary efficacy variable. The hazard ratio may be used for purposes of statistical hypothesis testing and as one indication of the amount of benefit (an increase in the odds of healing), but other measures must also be applied to understand the full importance of the study. Useful parameters on the time scale include the mean and median times as well as other percentiles to the study endpoint across treatment groups, and the median ratio.

Measuring effects on the time scale is particularly useful when the event of interest eventually occurs in 100% of the population (the time-to-event curve drops to zero), such that risk at the end of follow-up is not an issue. This is commonly the case in acute disease therapy with illnesses such as herpes labialis, herpes genitalis, herpes zoster, the common cold, otitis media, community-acquired pneumonia, and cellulitis in immunocompetent hosts.

Time effects are sometimes erroneously inferred from the hazard ratio. The use of terms such as accelerated time, re-

solved faster, twice as fast, and time was shorter in the context of the hazard ratio value suggests to the reader that magnitudes of change equivalent to the hazard ratio have been effected on the time scale, when in fact it is risk that has been changed. Changes in risk and time can be considerably different.

REFERENCES

1. **Allison, P. A.** 1995. Survival analysis using the SAS system: a practical guide. SAS Institute, Inc., Cary, N.C.
2. **Cox, D. R.** 1972. Regression models and life tables. J. R. Stat. Soc. B **34:** 187–220.
3. **Cox, D. R., and D. Oakes.** 2001. Analysis of survival data. Chapman and Hall, London, England.
4. **Harrell, F. E., Jr., R. M. Califf, D. B. Pryor, K. L. Lee, and R. A. Rosati.** 1982. Evaluating the yield of medical tests. JAMA **247:**2543–2546.
5. **Harrell, F. E., Jr., K. L. Lee, R. M. Califf, D. B. Pryor, and R. A. Rosati.** 1984. Regression modelling strategies for improved prognostic function. Stat. Med. **3:**143–152.
6. **Huff, J. C., A. Drucker, A. Clemmer, O. L. Laskin, Y. J. Bryson, and H. H. Balfour, Jr.** 2001. Effect of oral acyclovir on pain resolution in herpes zoster:a reanalysis. J. Med. Virol. **1:**93–96.
7. **Katz, M. H.** 1999. Multivariate analysis: a practical guide for clinicians. Cambridge University Press, New York, N.Y.
8. **Spruance, S. L., S. K. Tyring, B. DeGregorio, C. Miller, and K. Beutner.** 1996. A large-scale, placebo-controlled, dose-ranging trial of peroral valaciclovir for episodic treatment of recurrent herpes genitalis. Valaciclovir HSV Study Group. Arch. Intern. Med. **156:**1729–1735.
9. **Tyring, S., R. A. Barbarash, J. E. Nahlik, A. Cunningham, J. Marley, M. Heng, T. Jones, T. Rea, R. Boon, R. Saltzman, and Collaborative Famciclovir Herpes Zoster Study Group.** 1995. Famciclovir for the treatment of acute herpes zoster: effects on acute disease and post herpetic neuralgia. Ann. Intern. Med. **123:**89–96.
10. **Tyring, S. K., J. M. Douglas, L. Corey, S. L. Spruance, J. Esmann, et al.** 1998. A randomized, placebo-controlled comparison of oval valacyclovir and acyclovir in immunocompetent patients with recurrent genital herpes infections. Arch. Dermatol. **134:**185–191.
11. **Wood, M. J., R. Kay, R. H. Dworkin, S. J. Soong, and R. J. Whitley.** 1996. Oral acyclovir therapy accelerates pain resolution in patients with herpes zoster: a meta-analysis of placebo-controlled trials. Clin. Infect. Dis. **22:**341–347.
12. **Whitley, R. J., H. Weiss, J. W. Gnann, Jr., S. Tyring, G. J. Mertz, P. G. Pappas, C. J. Schleupner, F. Hayden, J. Wolf, S. J. Soong, et al.** 1996. Acyclovir with and without prednisone for the treatment of herpes zoster: a randomized, placebo-controlled trial. Ann. Intern Med. **125:**376–383.