Nonparametric estimator of relative time with application to the Acyclovir Prevention Trial

Stephen R Cole, Haitao Chu and Lei Nie

Background  Relative hazard is a central measure of association in randomized clinical trials. Relative time (RT) is a competing measure that is rarely used.

Purpose  We describe a simple area-based nonparametric estimator of RT and illustrate its use in the Acyclovir Prevention Trial.

Methods  Let \( Q_x(p) \) be the quantile function for the \( x \)th treatment group, defined as the time by which \( p\% \) of the treatment group experience the event, and \( p_x \) be the maximum event proportion observed. Our consistent estimator is defined as the ratio of the integrals of \( Q_1(p) \) and \( Q_0(p) \) with integration over 0 to \( p \), where \( p = \min(p_1, p_0) \). Confidence limits (CL) are provided by bootstrap.

Results  A total of 703 immunocompetent adult men and women (54% male, 79% Caucasian, median age 49 years) with a history of ocular herpes simplex virus (HSV) were enrolled in 1992–1996, randomized to acyclovir or placebo, followed for up to 1 year for the 1st episode of ocular HSV, and 170 events were confirmed by a study-certified ophthalmologist using slit-lamp biomicroscopy. The nonparametric RT comparing acyclovir use with nonuse was 2.6 (bootstrap 95% CL: 1.6, 4.2). For comparison, the best-fitting parametric model was the lognormal (RT = 2.5; 95% CL: 1.5, 3.9). In limited simulations, the average proposed estimate of RT was similar to the true RT with a relative root mean squared error of 1.13 compared to a correctly specified parametric (lognormal) model.

Limitations  An analytical variance estimator for the proposed RT is lacking. Also, more examples and more extensive simulations are warranted.

Conclusions  Similar to Cox’s relative hazard estimator, the proposed RT does not assume the data are generated from a particular distribution. RTs should be more widely used as a measure of association in clinical trials. Clinical Trials 2009; 6: 320–328. http://ctj.sagepub.com

Introduction

Relative hazards, typically estimated by Cox proportion hazards regression models [1], are a central measure of association for \( k \)-group comparisons in randomized clinical trials. The hazard at time \( t \) is the limit, as \( \Delta t \to 0^+ \), of the conditional probability of endpoint occurrence between \( t \) and \( t + \Delta t \) given that the endpoint had not occurred by \( t \) divided by \( \Delta t \). The relative hazard is a measure of the relative increase or decrease in the hazard of a treatment group compared to a placebo group. Heuristically, the relative hazard can be viewed as a measure of the vertical distance between the two treatment-specific survival functions (or more accurately, the log relative hazard is the vertical distance between the log-log of the survival functions). A deep understanding of the hazard requires knowledge of limits and conditional probability. Therefore some concepts may be difficult to convey to clinical and lay audiences. Indeed, Moser and McCann describe [2] an example where ‘... researchers had difficulty translating the concept of a hazard ratio into a clinically interesting difference between the
two treatment regimens.’ Clinical and lay audiences may use the hazard interchangeably with the risk, which is the probability of an endpoint in a group over a defined time period. This analogy may be problematic, as for example, the hazard is intrinsically susceptible to selection bias [3–6], while the risk is not.

Relative times (RTs), typically estimated by parametric survival models [7–9], are a competing measure of association for k-group comparisons in randomized clinical trials. The survival time at quantile or proportion $p$ is the time at which $p$ percent have developed the endpoint. The RT is a measure of the expansion or contraction in the percentiles of time for a treatment group compared to the placebo group. In contrast to the relative hazard, the RT is a measure of the horizontal distance between the two treatment-specific survival functions. Semiparametric approaches to estimate the RT have been developed [10–12], but rarely implemented in biomedical research. Communication of RTs to clinical and lay audiences only requires understanding of Figure 1(a), which is straightforward, irrespective of the methods used to obtain estimates of the RT.

First, we describe a standard parametric approach to estimating RTs and define a novel nonparametric estimator of RT that is based on the area to the left of treatment-specific survival functions. Second, we describe the Acyclovir Prevention Trial [13]. Third, we present results of applying the proposed method, as well as a series of parametric models to the Acyclovir Prevention Trial data. Fourth, we present the results of limited simulations. Last, we close with a discussion of strengths, limitations, and opportunities for extending the proposed method.

**Methods**

A standard approach to estimation of RTs is to employ parametric models fit by maximum likelihood. Say the observed data consist of $N$ copies of the independent and identically distributed triple $(T_i, \Delta_i, X_i)$. Let $T_i = \min(T_i^*, D_i, C_i)$, where $T_i^*, D_i$, and $C_i$ are random variables denoting the time from randomization to the study endpoint, loss to follow up, and administrative right-hand censoring, respectively, for patient $i$, where $i = 1 \text{ to } N$. Let $\Delta_i = 1$ if an endpoint occurred during follow up, i.e., $T_i^* < \min(D_i, C_i)$, and let $X_i = 1$ if patient $i$ was randomized to the treatment group. As is typical, we assume both loss to follow up and administrative censoring are independent of the endpoint process, or formally $f(T_i^*) = f(T_i^*|D_i, C_i)$, where $f(\cdot)$ is the density function.

In particular, a generalized gamma (GG) formulation [14] of the accelerated failure time model [15] is

$$\log T_i = \beta_0 + \beta_1 X_i + \sigma \log \varepsilon_i,$$

where $\beta_0$ is an intercept, $\beta_1$ maps the association between the treatment indicator $X_i$ and the time-to-event as the log RT, $\sigma$ is a scale parameter, and $\varepsilon_i$ is an error term with probability density function

$$f(\varepsilon) = \frac{|\lambda|^{\alpha/2} \exp(-\varepsilon^2/\lambda^2)}{\varepsilon \times \Gamma(1/\lambda^2)}, \quad (1)$$

where $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$. In the above, as $\lambda \to 0$, Equation (1) converges to a lognormal distribution, $\lambda = 1$ corresponds to a Weibull distribution, $\sigma = 1$ corresponds to a gamma distribution, and $\lambda = \sigma = 1$ corresponds to an exponential distribution.

![Figure 1](http://ctj.sagepub.com)  
**Figure 1** Illustration of RT by areas (1a) and ratio of restricted means (1b)
Informally, one can think of a nonparametric estimate of relative time (RT') as the treatment quantile area divided by the placebo quantile area. The treatment (or placebo) quantile area is defined as the vertical integration of the quantile function from the 100% survival down to the (100-p)% survival for the treatment (or placebo) group. This idea is depicted in Figure 1(a). Let \( Q_x(p_x) \) be a quantile from the observed data, where \( p_x \) is the percent of the group \( x \) that incurred the endpoint at the end of follow up. Because \( p_x \) is the endpoint proportion at the end of follow up, \( Q_x(p_x) \) is the maximum quantile estimable from the observed data. Then our nonparametric area-based RT is defined as

\[
RT' = \frac{\int_0^\pi Q_1(p) dp}{\int_0^\pi Q_0(p) dp},
\]

where \( \pi \) is the minimum of \( p_0 \) and \( p_1 \). Note that the quantile areas in Figure 1(a) are not equal to the restricted means, unless \( \pi \) goes to 1 as illustrated in Figure 1(b). A formal definition of our proposed non-parametric estimator of RT is given in Appendix 1. A proof of the consistency of this estimator is given in Appendix 2. Simply put, the time for group \( x \) is taken as the area below the survival function but above a horizontal line drawn at the low-point (i.e., \( (100-p_x)% \)) of the survival curve. When comparing two (or more) survival curves, we draw the horizontal line at the highest low-point among the curves, or at the minimum \( p_x \), which we call \( \pi \). With such an area-based time calculated for each group, we simply divide the times to obtain a RT. A pleasant feature of this metric is that \( RT' = E(T_1)/E(T_0) \) as \( \pi \to 1 \), or in words our estimator is a measure of the ratio of expected survival times comparing the treated group versus the placebo group as the event rates go to 1. Perhaps an unpleasant feature of this metric is that when (as typical) \( \pi < 1 \), and the RTs are nonproportional, the observed RT' is a function of the length of study follow up. While unpleasant, this dependence on study length holds equally for the standard parametric RT described above. When the RTs are proportional \( \pi < 1 \) only has implications for precision. Finally, to create a measure of precision for this proposed nonparametric estimate of RT, we employ nonparametric bootstrap [16] confidence limits (CL).

**Acyclovir prevention trial**

Between September 1992 and December 1996, the Herpetic Eye Disease Study Group randomized 703 immunocompetent men and women to the Acyclovir Prevention Trial [13]. Eligible patients were over 11 years of age and had a medically documented episode of ocular herpes simplex virus (HSV) recurrence in the prior 12 months. Eligible patients were excluded if they were on antiviral or immunosuppressive therapy, or had a history of immune dysfunction. Patients were randomized with equal allocation and a permuted-block design to either 400 mg of oral acyclovir taken twice daily, or placebo. Patients and clinical personnel were masked to the treatment assignment.

Patients were seen by a study-certified ophthalmologist at the onset of symptoms or at planned study visits at 1, 3, 6, 9, and 12 months post-randomization. The primary study endpoint was first documented recurrence of ocular HSV, assessed by a study-certified ophthalmologist using slit-lamp biomicroscopy. Time was counted in days from randomization to the minimum of: a documented study endpoint, loss to follow up, or administrative censoring at 365 days of follow up.

The study was overseen by an independent data and safety monitoring board, institutional review boards approved the protocol and informed consent forms, and each patient gave written informed consent. Further details of the Acyclovir Prevention Trial are available from the National Technical Information Service [17].

**Results**

Characteristics of the 703 men and women enrolled in the Acyclovir Prevention Trial are presented in Table 1 by randomized group, as well as overall. Briefly, at randomization the overall median age was 49 years, 54% were male, and 79% were Caucasian. As expected, none of these characteristics differed meaningfully by randomized group.

Of the 703 participants, 170 incurred a study endpoint during follow up, 48 or 7% were lost to follow up, and 485 completed study follow up free of an endpoint at 365 days. The endpoint rate was 6 recurrences per 10^4 person-days for the acyclovir group and 11.1 recurrences per 10^4 person-days for the placebo group. Recurrence-free treatment-specific survival functions are

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Acyclovir</th>
<th>Placebo</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=N=357</td>
<td></td>
<td>N=346</td>
<td>N=703</td>
</tr>
<tr>
<td>Age, years(^1)</td>
<td>49 (36, 66)</td>
<td>49 (34, 62)</td>
<td>49 (35, 64)</td>
</tr>
<tr>
<td>Male sex, % (n)</td>
<td>55% (197)</td>
<td>52% (180)</td>
<td>54% (377)</td>
</tr>
<tr>
<td>Caucasian, % (n)</td>
<td>80% (284)</td>
<td>78% (269)</td>
<td>79% (553)</td>
</tr>
</tbody>
</table>

\(^1\)median (interquartile range).
illustrated in Figure 2, with a parametric lognormal fit overlaid.

The nonparametric estimate of RT was 2.63, with bootstrap 95% CL: 1.58, 4.16 based on 5000 bootstrap replicates. Standard RT from a series of parametric accelerated failure time models are also presented in Table 2. The best fitting parametric model was the lognormal, regardless of whether one used the Akaike information criterion (AIC) or a nested log likelihood goodness-of-fit $\chi^2$ test statistic as the criterion. The lognormal model provided a RT $= 2.46$ (95% CL: 1.54, 3.94). The GG model provided a similar RT as the lognormal, but with a 1.03-fold larger standard error due to the unnecessary extra parameter. The Weibull model provided a 0.91-fold smaller RT compared to the lognormal and a 0.92-fold smaller standard error, but had AIC 10.9 units higher than the lognormal. The poorly fitting exponential model provided a 0.75-fold smaller RT compared to the lognormal, but with a 0.66-fold smaller standard error. For reference, the relative hazard was 0.55 (95% CL: 0.41, 0.75).

Figure 3 depicts the percentile-specific contributions to the estimates of the nonparametric RT. One of the two dimensions of the areas can be seen as the width of the ‘buildings’ in the solid cityscape: the other dimension is a increasing weight as one moves from left to right through the cityscape, but this is not depicted in Figure 3. Due to the heavier weighting of latter percentile-specific RTs, we believe a single RT appropriately summarizes these data, especially below the 94th percentile.

Because of the apparent nonconstancy of the RT in Figure 3, we further explored the assumption of proportional RTs in the Acyclovir Prevention Trial data. We fit series of more general parametric GG models which allow the shape and scale parameters to differ by treatment group. This exploration led to a lognormal model with a single scale parameter, as presented above. For instance, the combined AIC for a generalized lognormal model (allowing non-constancy in the RT by the scale differing across treatment arms) was 1215.5; the more restricted standard combined lognormal model (enforcing constancy in the RT with a single scale parameter) provided a similar AIC of 1214.0 (Table 2). A test of the difference between these two nested models yielded $\chi^2 = 0.485$ on 1 degree of freedom and a $P$-value of 0.486. Therefore, the data suggest that the proportional times assumption implied by the more restricted lognormal model was justifiable in these data.

### Simulations

We conducted a limited Monte Carlo simulation study to demonstrate the simulated bias and variance of our proposed nonparametric estimator and a (miss-specified) exponential model both compared to a (correctly specified) lognormal model. Simulations were generated to replicate the general features of the Acyclovir Prevention Trial, as well as a few other scenarios. Lognormal survival times were generated for 10,000 samples of 350 treated ($X = 1$) and 350 untreated ($X = 0$) patients using the model

$$
\log T_i = 7.0 + 0.9 \times X_i + 2.5 \times \log e_i
$$

where $\hat{\beta}_0 = 7.0142, \hat{\beta}_1 = 0.9011, \hat{\sigma} = 2.3653, \lambda \equiv 0$ were the maximum likelihood estimates of the intercept, logRT for treatment $X_0$, and scale

<table>
<thead>
<tr>
<th>Model</th>
<th>RT</th>
<th>95% CL</th>
<th>SE log RT</th>
<th>AIC^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonparametric</td>
<td>2.63</td>
<td>1.58, 4.16</td>
<td>0.244</td>
<td>NA</td>
</tr>
<tr>
<td>Parametric</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exponential</td>
<td>1.85</td>
<td>1.36, 2.51</td>
<td>0.157</td>
<td>1242.9</td>
</tr>
<tr>
<td>Weibull</td>
<td>2.25</td>
<td>1.46, 3.47</td>
<td>0.221</td>
<td>1224.9</td>
</tr>
<tr>
<td>Lognormal</td>
<td>2.46</td>
<td>1.54, 3.94</td>
<td>0.240</td>
<td>1214.0</td>
</tr>
<tr>
<td>General gamma</td>
<td>2.47</td>
<td>1.28, 4.02</td>
<td>0.249</td>
<td>1214.3</td>
</tr>
</tbody>
</table>

^1Based on a nonparametric bootstrap for nonparametric estimate.

^2$-2 \times \log \text{likelihood} + 2 \times \text{number of parameters}.$
parameter from the lognormal fit to the Acyclovir Prevention Trial data, and $e_i$ is an error term with a lognormal probability density function described in Equation (1). The 7% drop out was ignored, but we retained approximately 69% administrative censoring at 365 days. In addition, we explored the null scenario where $\beta_1 = 0.0$ and a smaller sample size of 75 treated and 75 untreated.

Table 3 presents a summary of results for the log RT obtained from a (correctly specified) lognormal model, a (miss-specified) exponential model, and our proposed nonparametric estimator. As expected, the average simulated estimates for the lognormal and nonparametric estimators are similar to the true log RT under both sample sizes and the alternative and null hypotheses of log RT $= 0.9$ and 0.0, respectively. Under the alternative hypothesis, the exponential model provided estimates of the log RT that were notably different from the true log RT and in the direction seen in the trial data. The standard deviation of the 10,000 estimates of the log RT provides a Monte Carlo estimate of the standard error. As expected, in all scenarios the proposed nonparametric estimator was less efficient than the correctly specified lognormal model: a price paid for protection against misspecification of the parametric model.

### Table 3  Results of 10,000 simulations of 700 or 150 patients with a true log RT of 0.9 or 0 due to equally allocated treatment based on a lognormal model

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Lognormal</th>
<th>Exponential</th>
<th>Nonparametric</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N = 700$, log RT = 0.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_1$</td>
<td>0.902</td>
<td>0.552</td>
<td>0.892</td>
</tr>
<tr>
<td>$\text{se}(\hat{\beta}_1)$</td>
<td>0.244</td>
<td>0.154</td>
<td>0.263</td>
</tr>
<tr>
<td>Root MSE$^1$</td>
<td>0.244</td>
<td>0.381</td>
<td>0.275</td>
</tr>
<tr>
<td>Relative root MSE</td>
<td>1</td>
<td>1.56</td>
<td>1.13</td>
</tr>
<tr>
<td>$N = 700$, log RT = 0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_1$</td>
<td>-0.002</td>
<td>-0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td>$\text{se}(\hat{\beta}_1)$</td>
<td>0.227</td>
<td>0.138</td>
<td>0.233</td>
</tr>
<tr>
<td>Root MSE$^1$</td>
<td>0.227</td>
<td>0.138</td>
<td>0.233</td>
</tr>
<tr>
<td>Relative root MSE</td>
<td>1</td>
<td>0.61</td>
<td>1.03</td>
</tr>
<tr>
<td>$N = 150$, log RT = 0.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_1$</td>
<td>0.907</td>
<td>0.563</td>
<td>0.877</td>
</tr>
<tr>
<td>$\text{se}(\hat{\beta}_1)$</td>
<td>0.536</td>
<td>0.343</td>
<td>0.570</td>
</tr>
<tr>
<td>Root MSE$^1$</td>
<td>0.536</td>
<td>0.481</td>
<td>0.570</td>
</tr>
<tr>
<td>Relative root MSE</td>
<td>1</td>
<td>0.90</td>
<td>1.06</td>
</tr>
<tr>
<td>$N = 150$, log RT = 0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\beta}_1$</td>
<td>-0.001</td>
<td>-0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td>$\text{se}(\hat{\beta}_1)$</td>
<td>0.498</td>
<td>0.305</td>
<td>0.507</td>
</tr>
<tr>
<td>Root MSE$^1$</td>
<td>0.498</td>
<td>0.305</td>
<td>0.507</td>
</tr>
<tr>
<td>Relative root MSE</td>
<td>1</td>
<td>0.61</td>
<td>1.02</td>
</tr>
</tbody>
</table>

$^1$Mean squared error, square root of squared bias plus variance.

Figure 3  RT by recurrence-free survival, dashed and dotted lines are the RT estimated by the proposed nonparametric method and the standard lognormal model, respectively. A horizontal reference line is drawn at the null RT of 1
In the smaller sample scenario, the difference between the estimated log RT from the exponential model and true log RT is less important than the imprecision and, therefore, the exponential model has a slightly lower MSE than the lognormal or nonparametric estimators (3rd panel). In the larger sample scenario, the difference between the log RT estimated by the exponential model and the true log RT outweighs any gain in precision by use of the exponential model as can be seen by the larger MSE compared to either the lognormal or nonparametric estimators (1st panel). In summary, the proposed nonparametric estimator of RT appeared to work well in the scenarios explored.

Discussion

Herein, we described a novel nonparametric estimator of RT, illustrated the method using data from the Acyclovir Prevention Trial, and compared results with standard parametric approaches to estimation of proportional RTs. We found that among adult men and women with herpetic eye disease, assignment to acyclovir expands their recurrence-free time by 2.63 that of placebo. The RT is a natural measure of association, perhaps more intuitive than the relative hazard. In our example, the nonparametric estimate of RT was 7% larger than the best fitting of the set of parametric models explored. The proposed nonparametric method does not assume that the data are generated from a particular survival distribution, and thereby should provide a robust estimate of the RT. When the data generation model and chosen parametric model exactly coincide the nonparametric and correct parametric approaches should both provide appropriate estimates of the RT, as demonstrated by the simulations. We hypothesize that exact coincidence of the data generation and analysis model to be rare in practice.

An estimate of the RT obtained with the proposed nonparametric method may have an associated cost in terms of a loss in precision. Using MSE as a criterion there should be cases where the nonparametric estimator is preferred to the best fitting parametric model and vice-versa. In our example, the nonparametric and lognormal estimates were almost exchangeable. Because the proposed nonparametric estimator is so easily calculated, we suggest a two-stage approach, whereby one simultaneously compares the estimated RT and its variance to determine the more appropriate estimator to emphasize in results. In such two-stage approaches it is only fair to present both the nonparametric and parametric estimates, or quantify the additional uncertainty due to the choice of estimator.

There are existing semiparametric approaches to estimating the RT [10–12]. However, computational issues have made these approaches diffuse slowly into the biomedical literature. In contrast we present a computationally simple approach that should be readily available to applied biostatisticians and practicing epidemiologists.

The relative hazard from a Cox proportional hazards model is widely used and accepted. However, the hazard may not be widely understood. It is possible that clinical and laboratory collaborators interpret the hazard ratio simply as a ratio of risks, rather than as a ratio of hazards. In many cases this interpretation is adequate, but not always. For example, the hazard (but not the risk) is intrinsically susceptible to selection bias [3–6].

One sometimes encounters the unfortunate practice of using the inverse of the relative hazard as a measure of RT. Recall the well known fact that the rate ratio is the inverse of the RT from an exponential survival distribution [18]. Further, note that with a nonconstant hazard as occurs with lognormal survival distributions, in the presence of considerable right-hand censoring (i.e., 76% in our example) the inverse of the relative hazard is approximately equal to the rate ratio (the approximate equality dissipates with diminished censoring). Using the Acyclovir Prevention Trial data, the inverse of the hazard ratio of 1.82 ̸= [2.63, 2.46] the nonparametric and best fitting parametric RT estimates, which reinforces the admonition against use of the inverse of the relative hazard as a measure of RT.

The proposed method can easily be extended in several directions. First, to allow for censoring to depend on measured covariates one could employ inverse probability-of-censoring weights [19]. In the present example, bias due to dependent censoring is an unlikely explanation for results due to low (i.e., 7%) loss to follow up. Second, to allow for the control of confounding as would be needed in nonrandomized studies one may employ inverse probability-of-treatment weights [20]. Inverse probability-of-treatment weights account for imbalances in measured covariates across levels of treatment while allowing the dimension of the data to remain the same as in the unadjusted case [21]. Third, while we present the two-group case one can extend the method presented here to more than two groups. Specifically, one could choose a reference group, estimate the survival function for each group, and then calculate the RT of each nonreference group compared to the reference.

There are several limitations to the present work. First, an analytic variance estimator is desirable, likely possible based on martingale theory for the
log RT, and the topic of future work. Such an analytic variance estimator will facilitate hypothesis tests, which may be needed in the clinical trial setting. Prior to development of such a variance estimator, one can obtain a nonparametric bootstrap based 1-or 2-sided 100(1-a)% confidence interval and a P-value for a given null hypothesis [16]. Second, the RT is a function of the length of follow up. However, rather than being seen as a limitation, one may view this as a natural consequence of studying the effect of treatments on the time to an endpoint. Third, more extensive Monte Carlo simulations are needed to (a) inform the choice of nonparametric versus parametric approach, and (b) compare the proposed nonparametric approach to existing semiparametric approaches. Fourth, more examples like the Acyclovir Prevention Trial are needed, preferably with survival functions that illustrate the relative merits of vertical (e.g., hazards) versus horizontal (e.g., times) contrasts.

For the data given in Figure 1, the nonparametric relative hazard (the vertical comparison) uses the data for all endpoints. However, the nonparametric RT (the horizontal comparison) only uses data on endpoints above a recurrence-free survival of about 0.8 and discards endpoints below 0.8 observed in the placebo group. One can imagine opposite scenarios where the RT uses all available endpoints, but the relative hazard does not. For example, if the placebo group were administratively censored at about 120 rather than 365 days, the nonparametric RT would use all of the available endpoints, but the nonparametric relative hazard would only use data up to 120 days, discarding endpoints in the treatment group observed beyond 120 days. Therefore, the observed survival functions inform which statistical approach utilizes a greater proportion of the observed endpoints. It is interesting to note that parametric versions of the RT or relative hazard use all the observed events (via extrapolation) irrespective of the structure of the survival functions.

We described a novel nonparametric estimator of RT, illustrated the method using data from the Acyclovir Prevention Trial, compared results with standard parametric approaches for estimation of RTs, and present limited supportive simulation results. In conclusion, RTs should be more widely used as a measure of association in clinical trials.

Acknowledgments

Dr Cole was supported in part by NIH, NIAID R03-AI 071763. Dr Chu is supported in part by the Lineberger Cancer Center Core Grant CA16086 from the NIH, NCI. This work was done while Dr Nie was working at Georgetown University. Views expressed in this article are the author’s professional opinions and do not necessarily represent the official positions of the US Food and Drug Administration.

The authors would like to thank Dr Roy Beck for his assistance with the Acyclovir Prevention Trial data, and Dr Alvaro Munoz for expert advice regarding the GG model.

References


Appendix 1. Formal definition of the nonparametric area-based RT estimator

Here we use the same notation given as in the Methods section. Below, we adopt usual conventions of representing realizations of random variables by lower-case letters and suppress the individual index $i$ where possible. Let $t_{x,1} < t_{x,2} < \cdots < t_{x,j} < \cdots < t_{x,m_x}$ be the $m_x$ distinct endpoint times for group $X = x$. For each $x$ and $j = 1$ to $m_x$, let $n_{x,j}$ and $d_{x,j}$ be the number of patients at risk just prior to $t_{x,j}$ and the number of patients with recurrence at $t_{x,j}$, respectively.

To calculate the nonparametric relative time $RT(x)$ given in Equation (2), first we determine the treatment-group specific survivor functions $\hat{S}_x(t_{x,j})$ using the Kaplan–Meier product limit estimator [22] as $\hat{S}_x(t_{x,j}) = \prod_{k=1}^{j-1} (1 - d_{x,k}/n_{x,k})$ or alternatively using the Nelson–Aalen estimator [23,24] as $\hat{S}_x(t_{x,j}) = \exp[-\sum_{k=1}^{j} d_{x,k}/n_{x,k}]$. Next, we define the proportion of individuals expected to fail at $t_{x,j}$ as $\hat{p}_{x,j} = \hat{S}_x(t_{x,j-1}) - \hat{S}_x(t_{x,j})$. Let $\hat{p}_x = \sum_{j=1}^{m_x} \hat{p}_{x,j}$ and $\hat{\pi} = \hat{p}_1/\hat{\pi}_0$. Then we calculate the area $A_x = \sum_{t_{x,j} > \pi} \hat{p}_{x,j} \times t_{x,j}$ as an approximated Darboux sum of the integral $\int_{0}^{\pi} Q_x(p)\,dp$.

We provide a proof of consistency, that $A_x \to \int_{0}^{\pi} Q_x(p)\,dp$ as $m_x \to \infty$, in Appendix 2. In other words, the area is calculated by summation of a series of horizontal rectangles, with coordinates: $(0,\hat{\pi}_{x,j-1})$, $(\hat{\pi}_{x,j-1}, t_{x,j})$, $(\hat{\pi}_{x,j}, 0)$, and $(0,\hat{\pi}_{x,j-1})$ when $\hat{\pi}_{x,j-1} > \hat{\pi}$. Finally, the nonparametric area-based relative time $RT(x)$ is then simply estimated by the ratio $A_x/\hat{\pi}_0$.

To contrast this formulation of relative time $RT(x)$ to the standard RT consider again a GG. Under a parametric GG, $Q_x(p) = e^{\hat{\beta}_x \cdot \hat{p}_x} \cdot \hat{p}_x^{P_x(p)}$, where $\hat{p}_x(p)$ is a standard GG$(0, 1, \lambda_x)$. Say $GG(\hat{\beta}_x, \sigma_x, \lambda_x)$ for exposure groups $x = 0, 1$. The RT at percentile $p$ is

$$RT(p) = e^{\hat{\beta}_1 \cdot \hat{p}_1} \times \{[q_x(p)]^{\sigma_x} / [q_y(p)]^{\sigma_y}\};$$

hence in this general form of GG the RT is not constrained to be constant over $p$. If $\lambda_1 = \lambda_0 = \lambda$, then the (still nonconstant) RT($p$) = $e^{\hat{\beta}_1 \cdot \hat{p}_1} \times [q_x(p)]^{\sigma_x - \sigma_0}$.

Further, if $\sigma_1 = \sigma_0 = \sigma$, then RT($p$) = RT = $e^{\hat{\beta}_1 \cdot \hat{p}_1}$. This last RT, which is constant over $p$, is what we term a standard parametric RT and make reference to below. The maximum likelihood estimate of RT($p$) is an asymptotically unbiased estimator assuming the failure times $T_i$ follow the accelerated failure time model given in Equation (1). A parametric analog to our RT at percentile $p$ is

$$RT'(p) = e^{\hat{\beta}_1 \cdot \hat{p}_1} \times \int_{0}^{\pi} \frac{[q_x(p)]^{\sigma_x}}{[q_y(p)]^{\sigma_y}} \cdot dp,$$

which is nonconstant over $p$ and does not generally equal RT as described in the Methods section. If $\lambda_1 = \lambda_0 = \lambda$, then the (still nonconstant) RT($p$) = $e^{\hat{\beta}_1 \cdot \hat{p}_1} \times (\int_{0}^{\pi} [q_x(p)]^{\sigma_x} \cdot dp) / (\int_{0}^{\pi} [q_y(p)]^{\sigma_y} \cdot dp)$ and still differs from the standard RT. Finally, further if $\sigma_1 = \sigma_0 = \sigma$ (the case we are interested in), then RT($p$) = RT = $e^{\hat{\beta}_1 \cdot \hat{p}_1}$. So, in the parametric setting, the proposed relative time RT’ is shown to equal the standard relative time RT under the assumption of proportional RTs.

Appendix 2. Consistency of the nonparametric area-based RT estimator

We shall first prove that $A_x = \sum_{\hat{\pi}_{x,j-1} \leq \hat{\pi}} (\hat{p}_{x,j} \times \hat{\pi}) - \hat{p}_{x,j-1} \times t_{x,j}$ converges to $\int_{0}^{\pi} Q_x(p)\,dp$ as the number of distinct event times $m_x \to \infty$. $A_x$ can be viewed as an approximated Darboux sum of the integral $\int_{0}^{\pi} Q_x(p)\,dp$.

$$A_x = \int_{0}^{\pi} Q_x(p)\,dp$$

$$= \sum_{\hat{\pi}_{x,j-1} \leq \hat{\pi}} (\hat{p}_{x,j} \times \hat{\pi}) - \hat{p}_{x,j-1} \times t_{x,j}$$

$$- \sum_{\hat{\pi}_{x,j-1} < \hat{\pi}} (\hat{p}_{x,j} \times \hat{\pi}) - \hat{p}_{x,j-1} \times t_{x,j}$$

$$+ \sum_{\hat{\pi}_{x,j-1} \leq \hat{\pi}} (\hat{p}_{x,j} \times \hat{\pi}) - \hat{p}_{x,j-1} \times t_{x,j}$$

$$- \sum_{\hat{\pi}_{x,j-1} < \hat{\pi}} (\hat{p}_{x,j} \times \hat{\pi}) - \hat{p}_{x,j-1} \times t_{x,j}$$

$$+ \sum_{\hat{\pi}_{x,j-1} < \hat{\pi}} (\hat{p}_{x,j} \times \hat{\pi}) - \hat{p}_{x,j-1} \times t_{x,j}$$

$$- \int_{0}^{\pi} Q_x(p)\,dp$$


Since $\hat{p}_{x,j-1}$ converges to $p_{x,j-1}$ weakly, see e.g., Theorem 6.3.1 [25], we can show that the 1st term $\sum_{\hat{p}_{x,j-1} \leq \hat{x}} [(\hat{p}_{x,j} \wedge \hat{x}) - \hat{p}_{x,j-1}] \times t_{x,j} - \sum_{\hat{p}_{x,j-1} \leq \hat{x}} ([p_{x,j} \wedge \pi) - p_{x,j-1}] \times t_{x,j}$ converges to 0 as $m_x \to \infty$, using Lebesgue's dominated convergence theorem [26]. It is obvious that the 2nd term $\sum_{\hat{p}_{x,j-1} \leq \hat{x}} [(p_{x,j} \wedge \pi) - p_{x,j-1}] \times t_{x,j} - \sum_{\hat{p}_{x,j-1} \leq \hat{x}} [(p_{x,j} \wedge \pi) - p_{x,j-1}] \times t_{x,j}$ also converges to 0 in probability as $m_x \to \infty$. Finally, the 3rd term $\sum_{\hat{p}_{x,j-1} \leq \hat{x}} [(p_{x,j} \wedge \pi) - p_{x,j-1}] \times t_{x,j} - \int_0^\infty Q_0(p)dp$ converges to 0 according to the definition of the integral. Therefore, $\lim_{m_x \to \infty} A_x = \int_0^\infty Q_0(p)dp$, which leads to $\lim_{m_x \to \infty} A_1/A_0 = (\int_0^\infty Q_1(p)dp)/(\int_0^\infty Q_0(p)dp)$, where $A_x = \sum_{\hat{p}_{x,j-1} \leq \hat{x}} [(\hat{p}_{x,j} \wedge \hat{x}) - \hat{p}_{x,j-1}] \times t_{x,j}$. 