TIME-DEPENDENT COVARIATES IN THE COX PROPORTIONAL-HAZARDS REGRESSION MODEL

Lloyd D. Fisher and D. Y. Lin
Department of Biostatistics, University of Washington, Seattle, Washington 98195-7232; e-mail: lfisher@biostat.washington.edu; danyu@biostat.washington.edu

KEY WORDS: survival analysis, longitudinal analysis, censored data, model checking

ABSTRACT

The Cox proportional-hazards regression model has achieved widespread use in the analysis of time-to-event data with censoring and covariates. The covariates may change their values over time. This article discusses the use of such time-dependent covariates, which offer additional opportunities but must be used with caution. The interrelationships between the outcome and variable over time can lead to bias unless the relationships are well understood. The form of a time-dependent covariate is much more complex than in Cox models with fixed (non–time-dependent) covariates. It involves constructing a function of time. Further, the model does not have some of the properties of the fixed-covariate model; it cannot usually be used to predict the survival (time-to-event) curve over time. The estimated probability of an event over time is not related to the hazard function in the usual fashion. An appendix summarizes the mathematics of time-dependent covariates.

INTRODUCTION

One of the areas of great methodologic advance in biostatistics has been the ability to handle censored time-to-event data. "Censored" means that some units of observation are observed for variable lengths of time but do not experience the event (or endpoint) under study. Such data were first studied and analyzed by actuaries. Kaplan & Meier presented the product limit or Kaplan-Meier curve to efficiently use all of the data to estimate the time-to-event curve (6).

Comparison of groups based on this nonparametric estimate is given by the log-rank test. Sir David Cox considered the introduction of predictor/explanatory variables or covariates into such models. The hazard may be thought of as proportional to the instantaneous probability of an event at a particular time. Cox (2) proposed a model in which the effect of the covariates is to multiply the hazard function by a function of the explanatory covariates. This means that two units of observation have a ratio of their hazards that is constant and depends on their covariate values. This model is usually called either the Cox regression model or the proportional-hazards regression model. It is important that covariates in this model may also be used in models in which the underlying survival curve has a fully parametric form, such as the Weibull distribution. At this time the Cox model is probably the most widely used model and is discussed in this paper, but the same issues, problems, and opportunities hold for the other parametric time-to-event models.

Because one of the most common uses of the model is for death as an end-point and also given the historical development in actuarial science, the field is sometimes referred to as survival analysis. In an industrial setting the events are often failure of devices, machines, etc, and the field is also referred to as failure time analysis.

The model is often used to examine the predictive value of, for example, survival, in terms of subject (often patients in some medical setting) covariates such as treatment, age, gender, height, weight, relative weight, smoking status, ethnicity categories, diastolic and systolic blood pressures, education, and income, to predict survival. The exponential of the coefficients from the Cox model gives the instantaneous relative risk for an increase of one unit for the covariate in question.

In many instances covariate data are collected longitudinally. For example, blood pressure, CD4 count, relative weight, disease history, and hospitalization data may be collected at selected periodic time points. As another example, treatment or other exposure may change over time. It seems natural and appropriate to use the covariate information that varies over time in an appropriate statistical model. One method of doing this is the time-dependent Cox or proportional-hazards model. This article discusses the use of such models. The article is written primarily for those who have a working familiarity with the usual fixed-covariate proportional-hazards model. The emphasis is on differences that arise when time-dependent covariates are used instead of fixed covariates. This review avoids mathematical details, which are relegated to an appendix for readers with more mathematical background.

We see below that the use of time-dependent covariates offers exciting opportunities for exploring associations and potentially causal mechanisms. Unfortunately we also see that the use of time-dependent covariates is technically

difficult in the choice of covariate form, has great potential for bias, and does not lead to prediction for the individual survival experience as does the usual Cox model with fixed covariate values.

Selecting the Form of a Time-Dependent Covariate

When a predictor or independent covariate is allowed to vary over time, one needs to determine its form over time. This introduces a number of subtleties and difficulties. Although this is easy to state, the application is more difficult. Several illustrations introduce some of the issues involved.

ILLUSTRATION 1: SMOKING AND SURVIVAL As an example of a time-dependent covariate, consider the effect of smoking on survival. Suppose that we have a cohort of individuals who are contacted at yearly intervals. As part of the interview process, data are collected on both current smoking status (defined as smoking any cigarettes during the prior month) and the estimated total number of cigarettes smoked over the past year. The hypothesis to be investigated is that current cigarette smoking increases the risk of death.

Perhaps the most immediate approach is to use a step-function that equals one if the individual was smoking at the last follow-up and zero if the person was not smoking at the last follow-up.

For example an individual who was alive after 4 years of follow-up and who was smoking at baseline, at years 1 and 2, but not at years 3 and 4, would have a time-dependent covariate that equals 1 up to the beginning of year 3 and then drops down to zero. Note the step when the smoking status changed. A step function is a function that takes on constant values on intervals.

Cavender et al (1) present an analysis by using time-dependent covariates. The data are from the Coronary Artery Surgery Study (CASS), in which individuals with mild angina (i.e. chest pain caused by coronary artery disease) were randomized to early coronary artery bypass graft surgery or early medical therapy. Data on smoking status were collected every 6 months, and, for the first analysis, a step function of the type described above (but with 6-months intervals) was used. Very surprisingly, the estimated effect of current cigarette smoking on survival was positive although not statistically significant. This led to an examination of the individual patient smoking histories. It turned out that most of those who had died were smokers, but many stopped smoking at the last follow-up before their death. In a number of instances this was apparently explainable by hospitalization for a myocardial infarction or congestive heart failure. One conjectures that other patients may have had prodromal symptoms prompting smoking cessation.

The investigators addressed the problem in two ways. One approach was to use a time-lagged covariate. At the first and second 6-month intervals,

the baseline smoking status was used. For subsequent intervals, the next to last follow-up value was used rather than the last assessment. This resulted in a statistically significant increased risk from cigarette smoking. A second approach (and that presented in this paper) used the percentage of the follow-up period that the subject smoked. This also resulted in a statistically significant detrimental association with smoking.

The point here is that the choice of a time-dependent covariate involves the choice of a functional form for the time-dependence of the covariate. This choice is usually not self-evident but may be suggested by biological understanding or biological hypothesis. Other examples illustrate this point.

ILLUSTRATION 2: CHOLESTEROL-LOWERING DRUGS Suppose that we have an observational database on many individuals with high lipid values who have been placed on cholesterol-lowering drugs. Suppose also that a number of risk factors are in the database and that we want to examine the effect of treatment with a cholesterol-lowering drug. First note that the usual fixed covariates can be used in a model along with time-dependent covariates. Thus baseline lipid values could be used (as well as a time-dependent cholesterol covariate) to allow adjustment for this and other baseline patient characteristics. (Formally one can think of a fixed covariate as a time-dependent covariate that happens to have a constant value for all time points. Although this is mathematically correct, computer software will run much faster if the covariates that do not change over time are entered as fixed covariates.) How might one model the introduction of a cholesterol-lowering drug at some time during follow-up? If the drug effect is only through the plasma lipids and if the lipid values are periodically collected during follow-up, then one might enter the lipid value from the last follow-up as the value at each time point. Suppose the primary manner in which the lipid values contribute to mortality is through the atherosclerotic process. Also assume that abnormally high levels lead to a lipid build-up in the vascular lesions and, by contrast, that very low levels lead to some leaching out of the lipids. What if we assume that both the build-up and leaching out take place over extended time periods? Then we might want to use some complex function over time to model the effect. For example risk might be modeled as a moving weighted average of the values over some long time interval. The weights might be highest for the current levels and then drop off as the values become more remote.

If the lipid values are not measured routinely but we know when lipid-lowering drug therapy is implemented and we assume that the effect will become more pronounced over time because of the mechanisms involved, then we might want a function that is zero until the therapy is introduced and then increases (e.g. linearly) over a fixed time interval (a few years?) to a value of one. Thus

rather than a step function that jumps up, the benefit is assumed to increase continually over a time interval before reaching the maximum effect.

ILLUSTRATION 3: CIRCADIAN VARIABILITY IN PLASMA LEVELS OF A PROTECTIVE DRUG Consider an antiplatelet drug to reduce complications of vascular disease, e.g. vascular death, myocardial infarction, and ischemic stroke, as a combination endpoint. The endpoint is reached at the first occurrence of any of the three components. The endpoints are assumed to be caused by the thrombotic process with platelets involved in producing the thrombosis. The protection is assumed to be directly related to the plasma level of the drug or biologic material being studied. Let us also assume that prior work with the compound given twice daily at a fixed dose has been reasonably used to model the plasma levels over time as a function of the time of dosing and selected patient covariate values, including weight, height, age, gender, and ethnicity, as well as measures from kidney and liver function tests. For each individual in the current study we can estimate the 24-hour circadian pattern of the plasma concentration (assuming the drug is taken at the times recommended). The model would be greatly enhanced if, at baseline, selected plasma values were taken for each patient to help in this modeling effort. The time of most of the events (except for death during sleep) is assumed to be recorded. How might we model the relationship with the drug (as mediated through the modeled plasma levels of the drug)? In addition to fixed patient covariates that increase the risk of an event (e.g. prior myocardial infarction, ischemic stroke, transient ischemic attack or peripheral arterial disease, or age), a time-dependent covariate that is the estimated plasma level at each time point for the individual patient could be used. Note that the production of this time-dependent covariate is quite complex. In the model there is a fixed parameter to be estimated that is the multiplier of the time-dependent covariate. If this is statistically significant, one might argue that there is a drug effect. However, it is well known that there is a circadian pattern to the time of occurrence of myocardial infarctions. The very early morning is a period of increased risk. To adjust for this, one could argue for circadian terms (for example the first five terms of the Fourier series) to adjust for the circadian pattern of events without appealing to a drug effect. In this case, for the drug to be effective it would need to add predictive power to the circadian terms standing alone.

On further reflection, the situation is even more complex. Why should the effect on the relative risk of an event be directly related to the plasma level of a drug based on the form of the Cox proportional-hazards model? Perhaps there is almost no drug effect until the plasma level reaches some value, and then the effect increases rapidly to a plateau at some other level. It may be that there is an intermediary measurement (e.g. in this case the percent reduction in platelet

aggregation as a function of plasma level) that would be more closely related and that could be modeled. Even when receptors are saturated, the drug effect is usually not perfect. Thus perhaps the model should begin at a plasma level of zero and increase to an asymptote at some finite multiple of the relative risk. All of these ideas can lead to different models. Further one could try a number of models; in this case the multiple-comparison issue could become severe (if there is no independent cohort to examine the performance of a final model). Although the chance to model is a strength of the time-dependent covariate approach, for exploratory data analyses it also is a weakness because trying too many models can lead to great overfitting of the data.

ILLUSTRATION 4: BONE MARROW TRANSPLANT GRAFTING Selected leukemias and other blood diseases are treated with bone marrow transplantation. In a bone marrow transplant (BMT), there is a conditioning regime to destroy the cancerous cells; after this conditioning, a transplant of bone marrow from the patient (autologous bone marrow), another person who has the same human leukocyte antigen haplotypes, or a matched unrelated donor is infused to the patient. Suppose that our endpoint of interest is survival. There is a substantial early mortality from the conditioning regimen, because there is a trade-off, with a stronger conditioning regimen increasing the risk of early death but reducing the risk of a relapse of the cancer. During approximately the first year, there is an increased risk of death from relapse of the original cancer. Further, both acutely and chronically there is a risk of the transplanted bone marrow attacking the recipient as foreign matter (graft versus host disease) and causing death. In the long term there are also the usual risks of death as well as an increased risk of other cancers from the treatment for the cancer at hand. How might we try to model this situation with time-dependent covariates?

If there are only bone transplant patients (who by definition live until the transplant is attempted), then the study cannot evaluate the effect of the BMT. Thus suppose we have a cohort of individuals diagnosed with a particular cancer that may or may not be treated with BMT at some stage. There may well be time-dependent covariates for other chemotherapeutic and radiologic treatments, as well as for the BMT. Here we discuss possible time-dependent covariates for the BMT.

The early risk of a death from the conditioning regime and the stress of the bone marrow transplant might be modeled by a step function that is one during the period of hospitalization and zero at other times. The risk of relapse over the first year after transplant could be modeled by a step function that is one over the first year after transplant and zero at other times. The transplant-related risk after year one could be modeled by a step function that equaled one from the end of the first year onward.

Prediction and Time-Dependent Covariates

One of the benefits of the proportional-hazards model with all fixed covariates is the ability to give individualized predictions of the estimated time to the event of interest. Estimated curves under different treatment modalities may be used for counseling patients. Estimated survival curves under smoking and nonsmoking scenarios can illustrate the added risk caused by smoking. Curves with different covariate values can illustrate the interaction between multiple risk factors. With time-dependent covariates the ability to predict is usually lost. One reason is that, because the model depends on the value of a changing quantity (the time-dependent covariate), at a future time the future values are usually unknown until they are actually observed. Second, if we know the future value of some covariates (e.g. blood pressure), the existence of the values implies that the subject has not reached a death endpoint. This fact implies that one cannot estimate the survival curve from the observed and future values of a quantity such as blood pressure. The existence of a positive value would imply that the subject was still alive. Knowing the covariate would imply knowledge at each stage of the vital status. (The appendix discusses these issues more, while discussing internal and external covariates.)

Interpretation and Time-Dependent Covariates

The time sequence of causal relationships is often difficult and tricky to untangle. For example, if a disease can cause wasting, then using weight as a time-dependent covariate to predict diagnosis may be somewhat of a self-fulfilling prophecy and, rather than predicting the disease, a lower weight may be an early sign of as yet undiagnosed disease.

Care must also be exercised when one assesses treatment efficacy with a regression model (1), which includes not only the treatment assignment but also a covariate (e.g. CD4 count and blood pressure) measured subsequent to the treatment assignment. If the effect of treatment on survival is predominantly mediated through the covariate, such an analysis will show little or no treatment effect on survival. This, however, can give useful information about the mechanism by which the treatment operates. In fact, this type of analysis has recently been used to assess the role of biological markers (such as CD4 count and viral load) as surrogate endpoints for clinical events (such as opportunistic infection and death) in HIV/AIDS trials (8). Certain covariates can be useful in reducing the bias of involving time-dependent exposure, as is discussed below.

In general if time-dependent covariates can change in relation to health or some other general concept related to the endpoint in the model, then interpretation is difficult and prone to be misleading. Great caution is advisable.

Discussion

The Cox proportional-hazards regression model for time-to-event data may be used with covariates, independent variables, or predictor variables that vary over time. These are called time-dependent covariates. Their use is much more complicated in practice than the fixed (time-independent) covariates. Further, the potential for erroneous inference and modeling is greatly increased. Still time-dependent covariates may be a powerful tool for exploring predictive relationships by using quantities that vary over time. The major points to remember include the following: (*a*) the modeling of a time-dependent covariate involves the choice of a function over time; this functional form may be far from obvious and may require deep biologic insight; (*b*) the choice of a complex functional form raises the possibility of too much modeling and great overfitting of a data set; (*c*) many time-dependent covariates are closely associated with the unit under study, are usually generated by that unit (e.g. blood pressure, weight, HIV status), and remove the usual relationship between the hazard function and the survival function; (*d*) time-dependent covariate models, except under certain circumstances, do not allow individual predictive time-to-event curves, which is different from the Cox model, with only fixed covariate values; (*e*) extreme caution must be exercised when modeling time-dependent exposure or treatment, especially if the change in exposure or treatment is related to the subject's health status; (*f*) most comprehensive statistical-software packages now provide proportional-hazards regression modeling with both fixed and time-dependent covariates allowed; (*g*) computer software exists for the estimation of coefficients and for examining the goodness-of-fit of the model.

The opportunities inherent in time-dependent modeling (including the explicit relationship of longitudinal values and the occurrence of an event) must be understood in light of the potential biases, strong assumptions needed in terms of the lack of other possible explanations, and the need to choose more complex functional forms for the modeling effort.

APPENDIX: MATHEMATICAL DISCUSSION

Time-Dependent Covariates and Proportional Hazards Regression

Let T be the failure time of interest, and let Z be a set of possibly time-dependent covariates. We use $Z(t)$ to denote the value of Z at time t, and $\bar{Z}(t) = \{Z(s) : 0 \leq s \leq t\}$ to denote the history of the covariates up to time t. It is convenient to formulate the effects of covariates on the failure time through the hazard function. The conditional-hazard function of T given $\bar{Z}$ is

$$\lambda(t \mid \bar{Z}) = \Pr(T \in [t, t + dt) \mid T \geq t, \bar{Z}(t),$$

where $(t, t + dt)$ is a small interval from t to $t + dt$. The Cox (2) proportional-hazards model specifies that

$$\lambda(t \mid \bar{Z}) = \lambda_0(t) e^{\beta' Z(t)}, \qquad 1.$$

where β is a set of unknown regression parameters and $\lambda_0(t)$ is an unspecified baseline hazard function.

Kalbfleisch & Prentice (5) distinguished between *external* and *internal* time-dependent covariates. This classification is helpful in interpreting the regression models and results for time-dependent covariates. An external covariate is one that is not directly related to the failure mechanism. One example would be the age of an individual in a long-term follow-up study. Another example would be the level of air pollution as a risk factor for asthma attacks. A third example would be times, for example, time of day or day of the year. On the other hand, an internal covariate is a value over time generated by the individual under study. Examples would include the Karnofsky score, blood pressures, procedural history, and CD4 counts measured over the course of the study.

A major difference between external and internal covariates lies in the relationship between the conditional hazard function $\lambda(t \mid \bar{Z})$ and the conditional survival function. The conditional survival function for a given covariate history is defined in general by

$$S(t \mid \bar{Z}) = \Pr(T > t \mid \bar{Z}(t).$$

For external covariates, this is also given by

$$S(t \mid \bar{Z}) = \exp\left(-\int_0^t \lambda(s \mid \bar{Z})\, ds \right),$$

which becomes

$$S(t \mid \bar{Z}) = \exp\left(-\int_0^t \lambda_0(s) e^{\beta' Z(s)}\, ds \right)$$

under Equation 1. By contrast, the conditional-hazard function bears no relationship to the conditional-survival function for internal covariates. In fact, the internal covariate requires the survival of the individual for its existence. For an internal covariate Z such as a Karnofsky score, $S(t \mid \bar{Z}) = 1$ provided that $Z(t)$ does not indicate that the individual has died. For the internal covariate of blood pressure, a measurable value indicates that the individual is still alive.

We now describe how to estimate the regression parameters of Equation 1. Suppose that we have n individuals in the study, such that the data consist of $\{X_i, \delta_i, \bar{Z}_i(X_i)\} i = 1, \ldots, n$, where X_i is the observation time (i.e. the last contact date) for the ith individual, and δ_i indicates, by the values 1 versus 0,

whether the ith subject fails or is censored at X_i, and $\bar{Z}_i(X_i)$ is the covariate history of the ith individual up to the observation time X_i. The estimation of β is based on the partial likelihood score function

$$U(\beta) = \sum_{i=1}^{n} \delta_i \left(Z_i(X_i) - \frac{\sum_{j \in R_i} e^{\beta' Z_j(X_i)} Z_j(X_i)}{\sum_{j \in R_i} e^{\beta' Z_j(X_i)}} \right), \qquad 2.$$

where R_i is the set of individuals who are at risk at X_i, that is, whose observation times are $\geq X_i$. The maximum partial likelihood estimator $\hat{\beta}$ is the solution to $U(\beta) = 0$. It is well-known that $\hat{\beta}$ is consistent and asymptotically normal with covariance matrix $I^{-1}(\hat{\beta})$, where

$$I(\beta) = \sum_{i=1}^{n} \delta_i \left(\frac{\sum_{j \in R_i} e^{\beta' Z_j(X_i)} Z_j(X_i) Z_j(X_i)'}{\sum_{j \in R_i} e^{\beta' Z_j(X_i)}} - \frac{\left\{ \sum_{j \in R_i} e^{\beta' Z_j(X_i)} Z_j(X_i) \right\} \left\{ \sum_{j \in R_i} e^{\beta' Z_j(X_i)} Z_j(X_i) \right\}'}{\sum_{j \in R_i} e^{\beta' Z_j(X_i)}} \right). \qquad 3.$$

It is apparent from Equations 2 and 3 that the statistical inference about β requires, at each uncensored X_i, the values of the covariates for all the subjects who are at risk at X_i. If Z varies its values continuously over time and is measured only at certain time intervals, then $Z_j(X_j)$ may not be available. In such situations, some interpolation between repeated measurements is necessary; see Lin et al (9) for a description of various interpolation schemes.

Epidemiologic cohort studies and disease prevention trials often involve follow-up on several thousand subjects for a number of years. The assembly of covariate histories, which requires biochemical analysis of blood samples or other specimens or the hand coding of individual diet records, can be prohibitively expensive if it is done on all cohort members. Under the case-cohort design (11), the covariate histories need only be assembled for all the cases (i.e. deaths) plus a random subset of the entire cohort. The aforementioned partial-likelihood method can be modified to analyze data from case-cohort studies (11, 12). Lin & Ying (10) proposed a general method for handling incomplete measurements of time-dependent covariates, including the case-cohort design as a special case.

Reducing Bias

One needs to be extremely cautious when interpreting results involving time-dependent exposure or treatment. A fundamental assumption in using models

like Equation 1 to formulate the effect of a time-dependent exposure or treatment on survival is that the change in exposure or treatment occurs in a random fashion. In many applications, however, individuals change exposure level or treatment for health-related reasons. Then the results based on a simple time-dependent exposure (treatment) indicator can be very misleading. Consider for example the simple scenario in which individuals will receive the treatment only when a certain intermediate event occurs. Suppose that the treatment is completely ineffective, but the occurrence of the intermediate event doubles the hazard of death. If the value of $Z(t)$ is 1 when the individual is on the treatment at time t and is 0 otherwise, then the hazard ratio parameter e^{β} in Equation 1 is equal to 2. This is the effect of the intermediate event, not the real effect of the treatment. Needless to say, such an analysis would be completely misleading.

It is possible to reduce the bias by properly adjusting for the factors that trigger the change in exposure level or treatment. Consider again the above scenario in which the treatment assignment is caused by the intermediate event, but now assume that not everyone who experiences the intermediate event will receive the treatment or that there are individuals who do not experience the intermediate event but will take the treatment. Then it is possible to fit models like

$$\lambda(t \mid Z) = \lambda_0(t)e^{\beta_1 Z_1(t)+\beta_2 Z_2(t)},$$

where $Z_1(t)$ indicates whether the individual is on the treatment at time t, and $Z_2(t)$ indicates whether the individual has experienced the intermediate event by time t. In this model, β_1 indeed pertains to the actual effect of the treatment (assuming that no other similar biases are in the model).

In practice, the reasons for changing exposure level or treatment may not be recorded or may be hard to quantify adequately. Then it is difficult to adjust for the factors that trigger the change through modeling. In such situations, modeling-time–dependent exposure or treatment should proceed with extreme caution and awareness of the potential biases involved.

Model Checking

Equation 1 assumes that covariates have proportionate effects on the hazard function over time. For instance, if $Z(t)$ is a single time-dependent indicator covariate, then Equation 1 implies that the hazard function is $\lambda_0(t)$ when $Z(t) = 0$ and is $\lambda_0(t)e^{\beta}$ when $Z(t) = 1$. It is important to assess this proportional-hazards assumption. There are a number of methods available in the literature. Here we discuss two methods that are most useful for time-dependent covariates.

Both methods are related to the score function $U(\beta)$ given in Equation 2. Lin et al (9) proposed the following class of weighted score functions:

$$U_w(\beta) = \sum_{i=1}^{n} W(X_i)\delta_i \left(Z_i(X_i) - \frac{\sum_{j \in R_i} e^{\beta' Z_j(X_i)} Z_j(X_i)}{\sum_{j \in R_i} e^{\beta' Z_j(X_i)}} \right),$$

where $W(t)$ is a weight function dependent on t. Let $\hat{\beta}_w$ be the solution to $U_w(\beta) = 0$. Suppose that $W(t)$ is a decreasing function of t and that the effects of Z are not proportional on the hazard function, but rather diminishing over time. In this case, $\hat{\beta}$ estimates an average of covariate effects over time, whereas $\hat{\beta}_w$ estimates a weighted average of covariate effects over time with more weights placed on the earlier part of the survival distribution, where the covariate effects are stronger. Therefore, $|\hat{\beta}_w|$ will tend to be larger than $|\hat{\beta}|$. This fact motivates us to test the proportional hazards assumption by comparing $\hat{\beta}$ and $\hat{\beta}_w$. Lin (7) derived a formal goodness-of-fit test based on this comparison.

The above test is sensitive to the choice of the weight function. More omnibus tests can be obtained by considering the following process:

$$U(\hat{\beta}; t) = \sum_{i: X_i \le t} \delta_i \left(Z_i(X_i) - \frac{\sum_{j \in R_i} e^{\beta' Z_j(X_i)} Z_j(X_i)}{\sum_{j \in R_i} e^{\hat{\beta}' Z_j(X_i)}} \right). \quad 4.$$

Note that $U(\hat{\beta}; t)$ is the partial likelihood score function based on the events that occur before time t. If the effects of covariates are indeed proportionate, then the solution to the equation $U(\beta; t) = 0$ should be similar to $\hat{\beta}$, regardless of the choice of t. In other words, $U(\hat{\beta}; t)$ should be close to 0 for all t. On the other hand, if the effects of covariates are not proportionate, then $U(\hat{\beta}; t)$ will tend to deviate from 0. Thus, it is reasonable to construct goodness-of-fit tests based on $U(\hat{\beta}; t)$. If Z is a single covariate, then the critical values for the supremum statistic $Q = \max_i I^{-1/2}(\hat{\beta})|U(\hat{\beta}; X_i)|$ are 1.628, 1.358, and 1.224 for the significance levels of 0.01, 0.05, and 0.10, respectively (13). If Z consists of multiple covariates, then the proportional-hazards assumption for the jth component of Z can be tested by the supremum statistic $Q_j = \max_i \{I^{-1}(\hat{\beta})_{jj}\}^{1/2}|U(\hat{\beta}; X_i)|$, where U_j is the jth component of U and $I^{-1}(\hat{\beta})_{jj}$ is the jth diagonal element of $I^{-1}(\hat{\beta})$. If the covariates are uncorrelated, then the critical values for Q_j are the same as those of Q described above (4). If the covariates are correlated, then the simulation technique of Lin et al (9) can be used to calculate the p values.

1. Cavender JB, Rogers WJ, Fisher LD, Gersh BJ, Coggin JC, Myers WO. 1992. Effects of smoking on survival and morbidity in patients randomized to medical or surgical therapy in the Coronary Artery Surgery Study (CASS): 10-year follow-up. *J. Am. Coll. Cardiol.* 20:287–94
2. Cox DR. 1972. Regression models and life tables. *J. Royal Stat. Soc. Ser. B* 34:187–220
3. Dickson ER, Grambsch PM, Fleming TR, Fisher LD, Langworthy A. 1989. Prognosis in primary biliary cirrhosis: model for decision making. *Hepatology* 10: 1–7
4. Fleming TR, Harrington DP. 1996. *Counting Processes and Survival Analysis*. New York: Wiley. 448 pp.
5. Kalbfleisch JD, Prentice RL. 1980. *The Statistical Analysis of Failure Time Data*. New York: Wiley. 321 pp.
6. Kaplan EL, Meier P. 1958. Nonparametric estimation for incomplete observations. *J. Am. Stat. Assoc.* 53:457–81
7. Lin DY. 1991. Goodness-of-fit analysis for the Cox regression model based on a class of parameter estimators. *J. Am. Stat. Assoc.* 86:725–28
8. Lin DY, Fischl MA, Schoenfeld DA. 1993. Evaluating the role of CD4-lymphocyte counts as surrogate endpoints in human immunodeficiency virus clinical trials. *Stat. Med.* 12:835–42
9. Lin DY, Wei LJ, Ying Z. 1993. Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* 80:573–81
10. Lin DY, Ying Z. 1993. Cox regression with incomplete covariate measurements. *J. Am. Stat. Assoc.* 86:1341–49
11. Prentice RL. 1986. A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika* 73:1–11
12. Self SG, Prentice RL. 1988. Asymptotic distribution theory and efficiency results for case-cohort studies. *Ann. Stat.* 16:64–81
13. Wei LJ. 1984. Testing goodness of fit for proportional hazards model with censored observations. *J. Am. Stat. Assoc.* 79:649–52