In: Survival Analysis, P.K. Andersen, N. Keiding, Section Editors; The Encyclopedia of Biostatistics, P. Armitage, T. Colton, Editors. Chichester, UK: John Wiley and Sons. 1998. pp. 4372-4389.

Structural Nested Failure Time Models

James M. Robins

Professor of Epidemiology and Biostatistics

Department of Epidemiology

Harvard School of Public Health

Boston, MA 02115

1. INTRODUCTION

Structural nested failure time models (SNFTMs) are causal models for the effect of a time-dependent treatment or exposure on a survival time outcome in the presence of time-dependent confounding covariates (Robins, 1989, 1992, 1993, 1995 1997; Robins et al., 1992; Witteman et al., 1997; Robins et al., 1994; Robins and Greenland, 1994; Mark and Robins, 1993ab). The simplest SNFTMs map a subject's observed failure time T, observed treatment and confounder history, and an unknown parameter ψ_0 into the time U that the subject would have failed if, possibly contrary to fact, treatment had been withheld. The causal parameter ψ_0 is identified if, as in a sequential randomized experiment, the treatment at time t is randomly assigned (i.e., ignorable) conditional on past treatment and confounder history. The method of g-estimation provides computationally convenient and robust semiparametric estimators of ψ_0 when ψ_0 is identified (Robins, 1992; Robins et al., 1992).

The usual approach to the estimation of the effect of a time-varying treatment on survival has been to model the hazard of failure at t as a function of past treatment history using a time-dependent proportional hazards model. In Section 2, we show that the usual approach may be biased, whether or not one further adjusts for past confounder history in the analysis, when *(a)* there exists a time-dependent risk factor for, or predictor of, the event of interest that also predicts subsequent treatment, and *(b)* past treatment history predicts subsequent risk factor level. The following two examples demonstrate conditions *(a)* and *(b)* will be true in studies in which either there is *(i)* treatment by "indication" and/or *(ii)* a time-dependent covariate that is simultaneously a confounder and an intermediate variable on the causal pathway from treatment to failure.

The drug AZT used in the treatment of AIDS is a direct red-blood cell toxin that is often withheld in anaemic subjects, since the toxic effects of AZT can worsen the anaemia. Further, anaemic patients are at increased risk of death. Thus in a study of the effect of AZT on survival of patients with AIDS, the time-dependent covariate anaemia is both a risk factor for death and a predictor of subsequent treatment with AZT. Further, as a red-blood-cell toxin, past AZT treatment is a risk factor for the development of anaemia. In occupational mortality studies, unhealthy workers who terminate employment early are at an increased risk of death compared to other workers and receive no further exposure to the chemical agent under study. Therefore, the time-dependent covariate, employment status at time t, is an independent risk factor for death and a predictor of future exposure to the study agent. In addition, previous exposure to the study agent may lead to early termination of employment if the agent causes a disabling illness. Epidemiologists refer to covariates such as anaemia or employment status in the above examples as time-dependent confounders.

The chapter is organized as follows. We first describe the fundamental assumption of no unmeasured confounders that, if true, allows us to test for and estimate causal effects from longitudinal data. In Section 2.3, we describe a valid α-level test, the g-test, of the null hypothesis of no causal effect of treatment on survival. In Section 2.4 - 2.5, we describe the potential for bias and lack of robustness of alternative testing procedures. In Section 3, we introduce the simplest SNFTMs — the deterministic SNFTMs. In Section 4, we show that g-estimation of deterministic SNFTMs provides a unified approach to estimation of and testing for the effect of a time-dependent treatment. In Section 5, we describe how the consequences of violations of our assumption of no unmeasured confounders can be explored through a sensitivity analysis. In Sections 2 - 5, we assume censoring is absent. In Section 6, we extend our methods to allow for censoring by end-of-follow-up, loss-to-follow-up, and competing risks. In Section 7, we show that it is difficult to incorporate *a priori* biological knowledge as restrictions on the functional form of our deterministic SNFTMs. However, it is straightforward to incorporate biological knowledge if we adopt a more general class of causal models, the instantaneous-rate rank-preserving structural nested failure time models (RPSNFTMs), which model the effect of a final instantaneous blip of treatment on survival. The parameters of a instantaneous-rate RPSNFTM can be consistently estimated using g-estimation. Instantaneous-rate RPSNFTMs allow the magnitude of the treatment effect to depend on the measured factors but not on unmeasured factors. This restriction is often biologically implausible. We therefore introduce, in Section 8, the class of instantaneous-rate SNFTMs which *(i)* allow the magnitude of the treatment effect to depend on both measured and unmeasured factors and *(ii)* include the instantaneous-rate RPSNFTMs as a special case. It is of scientific and public health interest to estimate the survival curves that would be expected under various treatment regimes in order to determine the optimal regime with which to treat future patients. Hence, we consider estimation of regime-specific survival curves. Finally we briefly describe an alternative non-nested structural failure time model that may sometimes have advantages in survival curve estimation.

Structural nested models for repeated measure and other non-failure time outcomes are considered in Robins (1989, 1993, 1994, 1997, 1998); they are not considered in this chapter.

2. CAUSAL INFERENCE from OBSERVATIONAL DATA

2.1. The Data

For pedagogic purposes, we shall consider a study of the effect of AZT treatment on the survival of AIDS patients. Let T_i be a continuous variable recording the survival time for the ith study subject, $i = 1, \ldots, n$, with time measured from study enrollment. Let $A_i(t)$ record subject i's AZT dosage rate at t and $L_i(t)$ record the value at t of a vector of various time-dependent and time-independent covariates such as CD4-lymphocyte count, presence of anaemia, and gender. For any time-dependent random variable $Z_i(t)$, let $\overline{Z}_i(t^-) = \{Z_i(u); 0 \le u < t\}$ be the history of the Z-process up to but not including time t and let $\overline{Z}_i(t)$ be the history of the process through t. Note that $Z_i(t)$ is defined only for $t \le T_i$. Until Section 6, we assume there is no censoring. In the absence of censoring, the observable variables are then $\left\{T_i, \overline{A}_i(T_i), \overline{L}_i(T_i)\right\}$, which we assume are independent and identically distributed, and henceforth suppress the i subscript denoting subject. Following Cox and Oakes (1984), Robins (1986, 1987), and Rubin

(1978), we shall also assume there exists a latent (possibly counterfactual) 'baseline' failure time random variable U representing a subject's survival time had, possibly contrary to fact, AZT always been withheld.

2.2. The fundamental assumption of no unmeasured confounders

Our fundamental assumption of no unmeasured confounders is

$$U \coprod A(t) \mid \overline{L}(t^{-}), \overline{A}(t^{-}), T \geq t, \tag{2.1}$$

where $A \coprod B \mid C$ means A is independent of B given C (Dawid, 1979). We will also refer to assumption (2.1) as the assumption that treatment $A(t)$ is sequentially ignorable or randomized given the past. Assumption (2.1) states that, conditional on AZT history and the history of all recorded covariates prior to t, increments in AZT dosage rate at t are independent of the baseline failure time random variable U. This assumption will be true if all risk factors for, i.e., predictors of, the baseline failure time U that are used by patients and physicians to determine the dosage of AZT at t are recorded in $\overline{L}(t^{-})$ and $\overline{A}(t^{-})$. For example, since physicians tend to withhold AZT from anaemic subjects, and in untreated subjects, anaemia is a predictor of survival, assumption (2.1) would be false if $\overline{L}(t^{-})$ does not contain anaemia history. It is the primary goal of the epidemiologists conducting an observational study to collect data on a sufficient number of covariates to ensure that our assumption (2.1) will be at least approximately true.

Assumption (2.1) is the fundamental condition that will allow us to draw causal inferences from observational data. It is precisely because (2.1) cannot be guaranteed to hold in an observational study and is not empirically testable that it is so very hazardous to draw causal inferences from observational data. Note that if, as in a sequentially randomized trial, at each time t, the dose of AZT was chosen at random by the flip of a coin, then (2.1) would be true even if the probability that the coin landed heads depended on past covariate and AZT-history. It is because physical randomization guarantees (2.1) that most people accept that valid causal inferences can be obtained from a randomized trial. See Rubin (1978), Robins (1986) and Holland (1986) for further discussion. In Section 5, we describe how the consequences of violations of (2.1) can be explored through sensitivity analysis.

For convenience, until Section 7, we shall assume that the treatment $A(t)$ received at time t is dichotomous, i.e., $A(t) = 1$ if on treatment at t and zero otherwise. Robins (1992) and Robins et al. (1992) consider non-dichotomous treatments. For $A(t)$ dichotomous, assumption (2.1) can be written

$$\lambda_A\left(t \mid \overline{A}(t^{-}), \overline{L}(t^{-}), U\right) = \lambda\left(t \mid \overline{A}(t^{-}), \overline{L}(t^{-})\right) \tag{2.2}$$

where, if $A(t)$ is a instantaneous-rate process, $\lambda_A\left(t \mid \overline{A}(t^{-}), \bullet\right) =$
$\lim_{\delta t \to 0} pr\left[A(t+\delta t) \neq A(t^{-}) \mid \overline{A}(t^{-}), T \geq t, \bullet\right] / \delta t$ is the hazard of the treatment process jumping

in the infinitesimal interval $[t, t + \delta t)$ given $\overline{A}(t^-)$ and $\bullet$. On the other hand, $\lambda_A\left(t \mid \overline{A}(t^-), \bullet\right) = pr\left\{A(t) \neq A(t^-) \mid \overline{A}(t^-), T \geq t, \bullet\right\}$ is a discrete hazard if $A(t)$ can only jump at non-random discrete times $t_1, t_2, \ldots$, as would be the case if the $A(t_k)$ recorded whether a subject was on AZT at weekly clinic visits. Due to measure theoretic subtleties, (2) but not (1) is a mathematically precise statement of our assumption of no unmeasured confounders.

2.3. A g-test of the causal null hypothesis

The sharp causal null hypothesis of no treatment effect on survival is that each subject's observed and baseline lifetimes are the same. That is,

$$U = T \quad \text{w.p.1} \tag{2.3a}$$

where w.p.1 stands for with probability 1. Given our assumption (2.2), the restriction on the distribution of the observables implied by (2.3a) is that the hazard of treatment jumps at t does not depend on the survival time T given past treatment and covariate history. That is,

$$\lambda_A\left(t \mid \overline{A}(t^-), \overline{L}(t^-), T\right) = \lambda_A\left(t \mid \overline{A}(t^-), \overline{L}(t^-)\right) \ . \tag{2.3b}$$

Hence, if the A process is a instantaneous-rate process, we can test (2.3b) by specifying a time-dependent Cox proportional hazards model

$$\lambda_0(t) \exp\left[\alpha' W(t)\right] \tag{2.4}$$

for $\lambda_A\left(t \mid \overline{A}(t^-), \overline{L}(t^-)\right)$, where $W(t)$ is a known vector valued function of $\left(\overline{A}(t^-), \overline{L}(t^-)\right)$, α is an unknown parameter vector, and $\lambda_0(t)$ is an unspecified baseline hazard function. If the A process jumps only at fixed discrete times, we interpret (2.4) as a model for the odds $\lambda_A\left(t \mid \overline{A}(t^-), \overline{L}(t^-)\right) / \left\{1 - \lambda_A\left(t \mid \overline{A}(t^-), \overline{L}(t^-)\right)\right\}$. If model (2.4) is correctly specified, an asymptotic α-level Cox partial likelihood score, Wald, or likelihood ratio test of the hypothesis $\theta = 0$ in the extended model that adds a term θT to $\alpha' W(t)$ in (2.4) is an asymptotically α-level test of the sharp null hypothesis (2.3a) under the assumption (2.2) of no unmeasured confounders. Robins (1992) refers to such a test as a g-test. Note a g-test first models *(i)* the hazard of the treatment process as a function of the survival time T and past treatment and covariate history and *(ii)* then tests whether the coefficient θ of T is significant. In fact, we obtain an α-level test of (2.3b) by testing $\theta = 0$ in the extended model that adds the term $\theta Q(t)$ to $\alpha' W(t)$ in (2.4) where $Q(t) = q\left(t, \overline{A}(t^-), \overline{L}(t^-), T\right)$ is a function chosen by the data analyst. The choice of $Q(t)$ effects the power but not the level of the g-test. The g-test is a generalization to time-dependent treatments and confounders of Rosenbaum's (1984) test for the effect of a single time-independent treatment.

2.4. Bias of standard methods

To understand why standard approaches that use Cox regression to model the hazard of failure as a function of past treatment history are biased whether or not one adjusts for past confounder history, we consider a group of AIDS patients who are alive at 10 months, dichotomized into those who developed anaemia by 8 months and those who remain free of anaemia at 8 months. A Cox regression analysis that estimates the rate ratio at 10 months attributable to AZT exposure in the interval $8-10$ months without adjusting for or stratifying on anaemia status can make AZT appear falsely beneficial, since anaemic subjects are at a higher risk of dying at 10 months and are less likely to receive AZT therapy in months $8-10$. That is, anaemia status at 8 months is a confounder for the causal effect of AZT treatment received in the interval from 8 to 10 months

However, because AZT causes anaemia, even if both AZT and anaemia have no causal effect on the survival of any subject, the above Cox regression analysis may continue to falsely suggest that AZT is beneficial even when we adjust for anaemia at 8 months. To see why, for simplicity, suppose now that 300 subjects receive AZT by 4 months, and 300 subjects never receive AZT. In both groups of 300, suppose that, regardless of AZT treatment or treatment for anaemia, 100 individuals are poor prognosis subjects who are destined to die at 10 months, 100 are moderate prognosis subjects destined to die at 20 months, and 100 are good prognosis subjects destined to die at 30 months. Suppose AZT causes anaemia in moderate prognosis patients. Specifically, all moderate prognosis patients would develop anaemia at 8 months if given AZT, whereas none would develop anaemia without AZT. All poor prognosis and no good prognosis patients develop anaemia regardless of AZT therapy. Under these assumptions, the data would be shown as in Table 1. Inspecting Table 1 , we observe that, within the stratum defined by the presence of anaemia at 8 months, the mortality rate at 10 months is less in those who received AZT than in those who did not. Similarly, in the stratum defined by the absence of anaemia, the mortality rate at 20 months is less in those who received AZT than in those who did not. Thus, a Cox analysis that adjusts for (or stratifies on) past anaemia history would falsely suggest that AZT has a beneficial effect on survival. This bias is attributable to the fact that in Table 1 *(a)* AZT by 4 months is a risk factor for subsequent anaemia, and *(b)* anaemia is a non-causal risk factor for death, since *(1)* the death rate at 10 months is greater in those with anaemia than in those without anaemia among subjects without AZT, and *(2)* the death rate at both 10 and 20 months is greater in those with anaemia than in those without anaemia among subjects with AZT. Note, by construction of our example, anaemia is not a causal risk factor for death; rather, it is a proxy for the unmeasured prognosis variable. Furthermore, anaemia is not an intermediate variable on the causal pathway from AZT treatment to death since, by construction, there is no such causal pathway.

Table 1. A Hypothetical Study

Anaemia by 8 Months	Time to Death (Months)			No Anaemia	Time to Death (Months)		
	10	20	30		10	20	30
No AZT	100*	0	0	No AZT	0	100†	100‡
AZT by 4 months	100*	100†	0	AZT by 4 months	0	0	100‡

* Poor prognosis patients.
† Moderate prognosis subjects.
‡ Good prognosis subjects.

It follows that we must control for the confounder "anaemia status at month 8" to estimate the causal effect of AZT in the interval $(8, 10)$. On the other hand, we must not control for the variable "anaemia status at month 8" to estimate the causal effect of AZT therapy in the interval $(0, 8)$ on survival. If, however, we summarize AZT history over the interval $(0, 10)$ in terms of cumulative dosage, average dose intensity, or the time since the initiation of AZT therapy, these requirements cannot be met, since we lose the ability to separate out AZT in the interval $(0, 8)$ from AZT in the interval $(8, 10)$. However, the g-test of Section 2.3 is specifically designed to control for confounding by variables affected by earlier treatment by never lumping treatment received at different times. Specifically, the g-test checks, at each time t, for association between treatment $A(t)$ received at t and the failure time T *(a)* after adjusting for confounder and treatment history before t, but *(b)* without adjusting for the "post-treatment" variables "covariate and treatment history subsequent to t." It is essential to the validity of the g-test that treatment history $\overline{A}(t^-)$ before t be adjusted for as a potential confounding factor for the effect of the treatment $A(t)$ received at t.

Formally, results in this section reflect the fact that the null hypothesis (2.3b) does not imply either that

$$\lambda_T\left(t \mid \overline{A}(t^-)\right) = \lambda_T(t) \tag{2.5}$$

or

$$\lambda_T\left(t \mid \overline{A}(t^-), \overline{L}(t^-)\right) = \lambda_T\left(t \mid \overline{L}(t^-)\right) \tag{2.6}$$

where $\lambda_T(t \mid \bullet)$ is the hazard of failure at t given $\bullet$. However, Robins (1986) proves that (2.3b) does imply (2.5) if either of the following are true:

$$\lambda_T\left(t \mid \overline{A}(t^-), \overline{L}(t^-)\right) = \lambda_T\left(t \mid \overline{A}(t^-)\right) \tag{2.7}$$

or

$$\lambda_A\left(t \mid \overline{A}\left(t^-\right), \overline{L}\left(t^-\right)\right) = \lambda_A\left[t \mid \overline{A}\left(t^-\right)\right] \ . \tag{2.8}$$

If (2.7) holds, we say $L(t)$ is not an independent predictor of failure. If (2.8) holds, we say $L(t)$ is not an independent predictor of subsequent treatment. If either (2.7) or (2.8) holds, we say that the $L(t)$ process is not a confounder for the effect of $A(t)$ on survival. In that case, we can test the sharp null hypothesis (2.3a) under Assumption (2.2) by ignoring data on $L(t)$ and testing (2.5) using a time-dependent Cox model for failure.

2.5. Computational complexity and non-robustness of tests based on the g-computation algorithm

In this Subsection, to avoid the need for product integral notation, we assume the covariate history $\overline{L}(T)$ can jump only at fixed times $k = 0, 1, 2, \ldots$. This is no practical limitation, since, for example, we could take the time interval between the jump times k and $k+1$ to be one second. The g-computation algorithm formula $r\left(t, \overline{a}, \overline{\ell}\left(m^-\right)\right)$ of Robins (1986, 1987) for the effect of a treatment regime $\overline{a} = \{a(u); 0 \leq u < \infty\}$ on survival to time t conditional on $\overline{L}\left(m^-\right) = \overline{\ell}\left(m^-\right)$ and $\overline{A}\left(m^-\right) = \overline{a}\left(m^-\right)$ is

$$r\left(t, \overline{a}, \overline{\ell}\left(m^-\right)\right) = \int \cdots \int \exp\left\{-\int_0^t \lambda_T\left(u \mid \overline{\ell}\left(u^-\right), \overline{a}\left(u^-\right)\right)\right\} \prod_{k=m}^{int(t)} dF\left[\ell(k) \mid \overline{\ell}\left(k^-\right), \overline{a}\left(k^-\right)\right] \tag{2.9}$$

where $int(t)$ is the greatest integer less than t, and $\lambda_T(u \mid \bullet)$ is the conditional hazard of failure at u given $\bullet$. Robins (1986, 1987) showed that under a sequential randomization assumption, $r\left(t, \overline{a}, \overline{\ell}\left(0^-\right)\right)$ is the probability of survival to t had, contrary to fact, all subjects followed treatment regime $\overline{a}$ until failure. Researchers studying causal models based on directed acyclic graphs (Spirtes, Glymour, and Scheines, 1993; Pearl, 1995) have recently rediscovered that g-computation algorithm formula. In addition, Arjas and Eerola (1993) and Klein et al. (1994) have also considered estimation of causal effects using this formula.

The g-null theorem of Robins (1986) states that (2.3b) is true if and only if, for all $\left(t, \overline{a}, \overline{\ell}\left(m^-\right)\right)$, $r\left(t, \overline{a}, \overline{\ell}\left(m^-\right)\right)$ depends on $\overline{a}$ only through $\overline{a}\left(m^-\right)$. It follows that one can, in principle, test the null hypothesis (2.3a) under the assumption (2.2) of no unmeasured confounders by *(i)* fitting a Cox proportional hazards model for $\lambda_T\left(u \mid \overline{a}(u), \overline{\ell}\left(u^-\right)\right)$ depending on parameters θ and a parametric or semiparametric model for $f\left(\ell(k) \mid \overline{\ell}\left(k^-\right), \overline{a}\left(k^-\right)\right)$ depending on parameters η, *(ii)* using the fitted model to construct an estimator $\widehat{r}\left(t, \overline{a}, \overline{\ell}\left(m^-\right)\right)$ of $r\left(t, \overline{a}, \overline{\ell}\left(m^-\right)\right)$ for various choices of $t, \overline{a}, \overline{\ell}(m^-)$ by evaluating the R.H.S. of (2.9), *(iii)* deriving estimates of the standard errors of the $\widehat{r}\left(t, \overline{a}, \overline{\ell}(m^-)\right)$, and finally *(iv)* constructing a test of the hypothesis that $r\left(t, \overline{a}, \overline{\ell}(m^-)\right)$ only depends on $\overline{a}$ through $\overline{a}(m^-)$ using the estimates $\widehat{r}\left(t, \overline{a}, \overline{\ell}(m^-)\right)$ and their estimated standard errors.

The difficulty with this procedure is two-fold. First, it is computationally extremely demanding since *(a)* the integral on the R.H.S. of (2.9) can not in general be evaluated analytically; a monte Carlo approximation must be used; and *(b)* it is difficult to compute delta-method estimators for the standard error of $\widehat{r}\left(t, \overline{a}, \overline{\ell}(m^-)\right)$ and bootstrap standard errors may be computationally too demanding because of *(a)*. Secondly, there will in general be no simple function ψ of the parameters (θ, η) that takes a fixed value (say zero) if and only if $r\left(t, \overline{a}, \overline{\ell}(m^-)\right)$ only depends on $\overline{a}(m^-)$. As discussed in Robins (1997), this fact implies that the tests based on the $\widehat{r}\left(t, \overline{a}, \overline{\ell}(m^-)\right)$ will be exquisitely sensitive to inevitable model misspecification. In summary, we suggest the g-test of Section 2.3 be used to test the null hypothesis (2.3b).

3. DETERMINISTIC STRUCTURAL NESTED FAILURE TIME MODELS

The g-test of the hypothesis $\theta = 0$ in the extension of model (2.4) is an asymptotic α-level test of the sharp null hypothesis (2.3a) if *(i)* model (2.4) is correctly specified, and *(ii)* the assumption (2.2) of no unmeasured confounders is true. However, we also wish to estimate the size of the treatment effect when the causal null is false. To do so, we introduce g-estimation of structural nested failure time models which will provide a unified approach to estimation of and testing for the effect of a time-dependent treatment.

The simplest SNFTM is a deterministic transformation model which assumes the counterfactual failure time U is a known function $h\left(T, \overline{A}(T), \overline{L}(T), \psi_0\right)$ of the subjects observed data $\left(T, \overline{A}(T), \overline{L}(T)\right)$ and an unknown parameter ψ_0; that is,

$$U = H(\psi_0), \tag{3.1a}$$

where

$$H(\psi) \equiv h\left(T, \overline{A}(T), \overline{L}(T), \psi\right) . \tag{3.1b}$$

A specific example of an deterministic SNFTM is the strong version of the accelerated failure time model of Cox and Oakes (1984) which assumes

$$h\left(T, \overline{A}(T), \overline{L}(T), \psi\right) = \int_0^T \exp\{\psi A(t)\}\, dt \ . \tag{3.2}$$

Any deterministic SNFTM (3.2) satisfies the following.

$$\text{If } \overline{A}(T) \equiv 0, \text{ then } U = T \ ; \tag{3.3a}$$

$$T \equiv U \text{ w. p. 1 if and only if } \psi_0 = 0 \ ; \tag{3.3b}$$

Eq. (3.3a) is a natural consistency assumption stating that if a subject is, in fact, untreated, then the observed failure time T equals the failure time U when treatment is withheld. Eq. (3.3b) implies the null hypothesis $\psi_0 = 0$ corresponds to the causal null hypothesis (2.3a) that treatment has no effect. To understand the implications of (3.2) when $\psi_0 \neq 0$, consider a subject who is continuously treated. Then, by (3.1)-(3.2), $U = e^{\psi_0} T$ so $T = e^{-\psi_0} U$. That is, a subject's untreated survival time U is expanded or contracted by the factor $e^{-\psi_0}$ by constant treatment. Hence, if $\psi_0 > 0$, treatment is harmful and lessens survival; if $\psi_0 < 0$, treatment is beneficial and increases survival. Robins et al. (1992) referred to deterministic SNFTMs as rank preserving structural failure time models.

4. g-ESTIMATION of ψ_0

We now describe how to obtain consistent asymptotically normal point and interval estimates of the parameter ψ_0 consistent with the g-tests of Section 2.3 in the sense that 95% confidence intervals for ψ_0 will fail to include zero if and only if the corresponding .05 level g-test rejects. As a simple example, consider the deterministic SNFTM (3.2) with ψ 1-dimensional. We estimate ψ by a "grid search." First, we note that for each value of ψ, $H(\psi)$ can be computed by (3.2) from the observed data. Hence, under the reasonable biological assumption that $|\psi_0| < 3$, separately, for each of the 61 values of ψ in the set $\{-3, -2.9, \ldots, 0, \ldots, 2.9, 3\}$, we perform a Cox partial likelihood score test (g-test) of the hypothesis $\theta = 0$ in the extended model that adds a term $\theta Q(t, \psi)$ to $\alpha' W(t)$ in (2.4) with $Q(t, \psi) = q\left(t, \overline{A}(t^-), \overline{L}(t^-), H(\psi)\right)$ a function chosen by the data analyst. A valid 95% large sample confidence interval for ψ_0 is the set of ψ for which the score test fails to reject at the .05 level provided our no-confounding assumption (2.2), our Cox model (2.4), and our deterministic SNFTM (3.1)-(3.2) are correct. Furthermore, the g-estimate $\widehat{\psi}$ is a consistent asymptotically normal estimator of ψ_0, where $\widehat{\psi}$ is defined to be the value of ψ

for which the partial likelihood score test of $\theta = 0$ is precisely zero. The parameter ψ is treated as a fixed constant when calculating the score test. The choice of the function $q()$ affects the length but not the coverage rate of the interval. The optimal choice of the function $q^*()$ is given in Robins (1992). The method of g-estimation can be extended to estimate the parameter, say, $\psi = (\psi_1, \psi_2)'$ of a multi-parameter deterministic SNFTM such as

$$H(\psi) = \int_0^T \exp\{\psi_1 A(t) + \psi_2 L^*(t) A(t)\} dt \ . \tag{4.1}$$

In model (4.1), $L^*(t)$ represents a known function of the covariate history $\overline{L}(t^-)$. If the true value ψ_{20} of ψ_2 is non-zero, there is a treatment-covariate interaction in the sense that the magnitude of the effect of the time-dependent treatment $A(t)$ depends on a subject's time-dependent covariate history $\overline{L}(t^-)$ through the function $L^*(t)$. A g-estimate of the parameter vector ψ of (4.1) is obtained by choosing $Q(t, \psi)$ to be a known vector-valued function of $\dim \psi$ chosen by the data analyst and θ to be a $\dim \psi$ valued parameter with $\dim \psi$ the dimension of the vector ψ.

4.1. Estimation with instrumental variables

Suppose $A(t) = (A_1(t), A_2(t))$ with $A_1(t)$ recording a physician's prescribed treatment and $A_2(t)$ recording the actual treatment at time t. One then might suppose that

$$A_1(t) \coprod U \mid \overline{L}(t^-), \overline{A}(t^-), T > t \tag{4.2}$$

is true but (2.1) is false if one believed that a predictor of both U and actual treatment $A_2(t)$ had not been included in $\overline{L}(t^-)$. Under Assumption (4.2), g-estimation of the parameter ψ_0 of the deterministic SNFTM (3.1) can proceed as before, except we view the model (2.4) as a model for the cause-specific hazard $\lambda_{A_1}\left(t \mid \overline{A}(t^-), \overline{L}(t^-)\right)$ for jumps in the $A_1(t)$ process, thus ignoring jumps in the actual treatment $A_2(t)$ process in our estimation procedure. In this setting, $A_1(t)$ is often referred to as an instrumental variable process, especially when prescribed treatment $A_1(t)$ has no direct causal effect on survival except through the actual treatment process $A_2(t)$. A familiar example of an instrumental variable is when $A_1(0)$ is the randomization indicator for assignment to treatment arm in a randomized clinical trial in which there is possibly non-random non-compliance and $A_2(t)$ is the actual treatment dose. (In such a setting, $A_1(t)$ can be defined to be zero by convention for times $t \neq 0$.) Then the g-estimation method described above is the method for adjusting for non-random non-compliance in randomized clinical trials described by Robins (1993, Appendix 1). For alternative rank estimation procedures, see Mark and Robins (1993b) and Robins and Tsiatis (1991).

5. SENSITIVITY ANALYSIS

In observational studies, our fundamental assumption (2.2) of no unmeasured confounders cannot be empirically tested from the data. Hence, it is important to conduct sensitivity analyses to determine how point and interval estimates for ψ_0 would change under increasingly severe violations of (2.2). Let η be a sensitivity parameter that we will vary (but not estimate) in our sensitivity analysis and consider the model

$$\lambda_A\left(t \mid \overline{A}\left(t^-\right), \overline{L}\left(t^-\right), U\right) = \lambda_0\left(t\right) \exp\left[\alpha' W\left(t\right) + \eta U\right] \ . \tag{5.1}$$

When $\eta = 0$, both the assumption (2.2) of no unmeasured confounders and our Cox model (2.4) are true. As $|\eta|$ increasingly deviates from zero, (2.2) is increasingly violated. Our goal, in a sensitivity analysis, is to obtain valid point and interval estimates for the causal parameter ψ_0 of our deterministic SNFTM under the assumption that (5.1) is correctly specified with α an unknown parameter to be estimated but with η known. Specifically, a 95% confidence interval for ψ_0 under these assumptions is obtained as the set of ψ for which the score test of the hypothesis $\theta = 0$ fails to reject at the .05 level in model

$$\lambda_A\left(t \mid \overline{A}\left(t^-\right), \overline{L}\left(t^-\right), H\left(\psi\right)\right) = \lambda_0\left(t\right) \exp\left[\alpha' W\left(t\right) + \eta H\left(\psi\right) + \theta Q\left(t, \psi\right)\right] \tag{5.2}$$

when η and ψ are treated as fixed and known when maximizing the partial likelihood over α. The data analyst should then display point and interval estimates for ψ for a moderately large number of choices for the sensitivity parameter η.

6. CENSORING

In this section, we extend our results to allow for right-censoring. We handle censoring by administrative end-of-follow-up differently from censoring by competing risks or by loss-to-follow-up.

6.1. Censoring by end-of-follow-up

We assume that there is a fixed known calendar date at which the follow-up of all subjects will end. We then define the potential censoring time C for a subject to be the difference between this end-of-follow-up date and the date at which the subject entered follow-up. Hence, the potential censoring time C is known for all subjects, even those who fail prior to the end-of-follow-up date. Because the potential censoring time C is known at start of follow-up ($t = 0$), we can and do regard C as a time-independent "pre-treatment" covariate that is contained in $\overline{L}(t^-)$ for each time $t \geq 0$. If, as we assume in this Section, the only cause of censoring is by end-of-follow-up, the data available for data analysis for each subject are $\left\{X = \min(T, C), \overline{A}(X), \overline{L}(X)\right\}$.

Since $H(\psi)$ can only be computed for uncensored individuals, it might seem natural when calculating g-estimates of ψ_0 to replace the now partially unobservable $H(\psi)$ by the new random variable $X^*(\psi)$ obtained by replacing T by X in (3.1b). Unfortunately, this approach fails since, if $\psi_0 \neq 0$, then $X^*(\psi_0)$ is not independent of $A(t)$ given $\left(\overline{A}(t^-), \overline{L}(t^-), X \geq t\right)$ even under the assumption (2.2) of no unmeasured confounders. Thus, an alternative approach is necessary. The key to our approach is to define new variables $(X(t,\psi), \Delta(t,\psi))$ that *(a)* in contrast to both T and $H(\psi)$, but like $X^*(\psi)$, are observed for all subjects, including those censored, and *(b)* like $H(\psi_0)$, but unlike $X^*(\psi_0)$, satisfy, for $t < C$,

$$\lambda_A\left[t \mid \overline{A}(t^-), \overline{L}(t^-), X(t,\psi_0), \Delta(t,\psi_0)\right] = \lambda_A\left[t \mid \overline{A}(t^-), \overline{L}(t^-)\right] \tag{6.1}$$

under assumption (2.2) and model (3.1). We can then estimate ψ_0 by g-estimation as before, except with $Q(t,\psi)$ now a function $q\left(t, \overline{L}(t^-), \overline{A}(t^-), X(t,\psi), \Delta(t,\psi)\right)$. Below, we define $X(t,\psi)$ and $\Delta(t,\psi)$ only for the deterministic SNFTM (3.2). Robins (1993, Appendix 4) gives the appropriate definitions for arbitrary SNFTMs. Let

$$X(t,\psi) = \min\{H(\psi), C(t,\psi)\}, \Delta(t,\psi) = I\{X(t,\psi) < C(t,\psi)\}$$

where $C(t,\psi) \equiv C - t + \int_0^t \exp\{\psi A(t)\} dt$ if $\psi \geq 0$ and $C(t,\psi) = \int_0^t \exp\{\psi A(t)\} dt + (C-t) e^{\psi}$ if $\psi < 0$. Eq. (6.1) is satisfied since $X(t,\psi_0)$ and $\Delta(t,\psi_0)$ are only functions of $\overline{A}(t^-), H(\psi_0)$ and C. Also $X(t,\psi)$ and $\Delta(t,\psi)$ are observables, since one can calculate that $X(t,\psi)$ is the minimum of the two observables $X^*(\psi)$ and $C(t,\psi)$. When $\Delta(t,\psi) = 0$, we say an individual is $\psi-$censored. Note that when $\psi \neq 0$, some failures will be $\psi-$censored. In practice, if efficiency is not of overriding concern, it is convenient to use a very simple function $Q(t,\psi)$ that produces reasonably efficient estimators of ψ_0. The indicator function $\Delta(t,\psi)$ has been found often to satisfy this criterion in a number of examples (Witteman et al., 1997). The efficient choice of $Q(t,\psi)$ is given in Robins (1993, Appendix 4). The simple sensitivity analysis methodology of Section 5 can be extended to the censored data setting by replacing ηU by $\eta X(\psi)$ in (5.2); however, better methodology should be developed.

6.2. Censoring by competing risks

In this section, we assume that in addition to censoring by end-of-follow-up C, there is additional censoring by loss-to-follow-up and/or competing risks. Let Q be the minimum of time to loss-to-follow-up or to a competing risk event. For ease of exposition, we no longer distinguish censoring by loss-to-follow-up from censoring by competing risks and simply refer to Q as time to censoring by competing risks. The data available are $X^* = \min(T,C,Q) = \min(X,Q)$, $\tau = I(X^* \neq Q)$, $\overline{A}(X^*)$, $\overline{L}(X^*)$ so that $\tau = 1$ if and only if a subject was either observed to fail or to reach end of follow-up without suffering a competing risk. To adjust for censoring by competing risks, we assume that we have recorded data on a sufficient number of potential confounding factors in $\overline{L}(t^-)$ so that there are no unmeasured confounders for censoring due to competing risk. That is,

$$\lambda_Q\left[t \mid \overline{A}(t^-), \overline{L}(t^-), X > t, X\right] = \lambda_Q\left[t \mid \overline{A}(t^-), \overline{L}(t^-), X > t\right] \tag{6.2}$$

in which case we shall also say that censoring by Q is ignorable given the past. Here, $\lambda_Q(t \mid \bullet)$ is the hazard for the random variable Q given $\bullet$.

Given the ignorable censoring assumption (6.2), our next task is to estimate the probability $K(X)$ of a subject surviving to $X = \min(T,C)$ without suffering a competing risk which will be used as an inverse weight in the weighted g-estimation procedure described below. To do so, we fit the Cox proportional hazard model

$$\lambda_{0Q}(t)\exp\{\alpha^{*\prime}W^*(t)\} \tag{6.3}$$

for the hazard $\lambda_Q\left(t \mid \overline{A}(t^-), \overline{L}(t^-), X > t\right)$ where $W^*(t)$ is a known vector valued function of $\overline{A}(t^-)$ and $\overline{L}(t^-)$, α^* is the vector of unknown parameters, and $\lambda_{0Q}(t)$ is an unspecified baseline hazard. We then estimate $K(X)$ by multiplying together the estimated conditional probabilities of not suffering a competing risk prior to X using the time-dependent Cox-model version of the

Kaplan-Meier estimator. Specifically, at each time Q_j where any subject j suffered a competing risk, we compute the Cox baseline hazard estimator

$$\widehat{\lambda}_Q(Q_j) = 1/\sum_{i=1}^{n}\left\{\exp\left[\widehat{\alpha}^{*\prime}W_i^*(Q_j)\right]I(X_i^* \geq Q_j)\right\}$$

of $\lambda_{0Q}(Q_j)$. We then estimate a subject's $K(X)$ by the Cox model version of the Kaplan-Meier estimator

$$\widehat{K}(X) = \prod_{\{j;Q_j \leq X, \tau_j = 0\}}\left\{1 - \widehat{\lambda}_Q(Q_j)\exp\left[\widehat{\alpha}^{*\prime}W^*(Q_j)\right]\right\}$$

which is the product, over the competing risk times $Q_j < X$, of the subject's estimated conditional probabilities of not suffering a competing risk. Note that a subject's estimated probability $\widehat{K}(X)$ depends on his/her treatment and covariate history through the covariate $W^*(t)$.

Having estimated $\widehat{K}(X)$ for each subject with $X = \min(T, C)$ observed, we then estimate ψ_0 by replacing, in our g-estimation procedure of Section 6.1, the function $Q(t,\psi)$ by $Q^*(t,\psi) \equiv Q(t,\psi)/\widehat{K}(X)$ for each person who did not suffer a competing risk ($\tau = 1$) and by $Q^*(t,\psi) = 0$ for each person who did ($\tau = 0$). For example, if we use the simple function $Q(t,\psi) = \Delta(t,\psi)$, then $Q^*(t,\psi) = \tau\Delta(t,\psi)/\widehat{K}(X)$.

We now give an intuitive explanation of why the g-estimate $\widehat{\psi}$ obtained by this method should be consistent for ψ_0. Given the correctness of our Cox model (6.3) and of our assumption of ignorable censoring by competing risks (6.2), the following will be true: for each person with X observed ($\tau = 1$) and an estimated cumulative probability of, say, $\widehat{K}(X) = .25$ of avoiding censoring by competing risks, there would, on average, have been three other persons (i.e., ghosts) who *(i)* were censored by competing risks before X ($\tau = 0$), and *(ii)* would have had a similar value of X and a similar covariate and treatment history up to X, had censoring by competing risks been prevented. We therefore assign this person with $\tau = 1$ and $\widehat{K}(X) = .25$ a weight of 4 in the g-estimation procedure by multiplying her covariate $\Delta(t,\psi)$ by the factor 4; she needs to count not only for herself but also for the three other similar subjects for whom X could not be observed due to censoring by competing risks and thus had $Q^*(t,\psi)$ set to zero.

This argument can be formalized to prove that the "competing risk" g-estimator $\widehat{\psi}$ is a CAN estimator of ψ_0. However, the previous method of obtaining confidence intervals and p-values is no longer valid because the contributions to the Cox partial likelihood score of the extended Cox model (2.4) for the treatment process are no longer uncorrelated for two distinct reasons. First, $K(X)$ and, therefore, the time-dependent covariate $\tau Q(t,\psi)/K(X)$ at time t depend on a subject's treatment and covariate history beyond t, disrupting the "martingale" structure of the Cox partial likelihood score. Secondly, the probability $K(X)$ of avoiding a competing risk is replaced by the estimate $\widehat{K}(X)$ which depends on all the data. However, if we fit the extended model (2.4) using a Cox proportional hazards program that computes the so-called "robust variance" (Wei et al., 1994), the resulting g-intervals and tests are guaranteed to be conservative, i.e., in large samples, nominal 95% confidence intervals are guaranteed to cover ψ_0 at least 95% of the time and .05 level g-tests are guaranteed to reject the null hypothesis $\psi_0 = 0$ when true no more than 5% of the time. If the conservative "robust variance" g-intervals are too long to distinguish important substantive alternatives, narrower intervals that cover ψ_0 95% of the time in large samples can be obtained using the formulae provided in Appendix 1.

Remark: Often it is reasonable to assume that censoring by end-to-follow-up C is also ignorable. To incorporate this assumption, we redefine Q to be the minimum of time to loss-to-follow-up, competing risks, or end-to-follow-up and replace C by a constant c^* which is slightly less than the maximal follow-up time $\max\{C_i; i = 1, \ldots, n\}$ so that $K(c^*)$ is bounded away from zero w.p.1. Then g-estimation and testing can proceed as above.

6.3. Estimation of direct effects

Suppose now we wish to estimate the direct effect of AZT $A(t)$ on survival when another treatment, say aerosolized pentamidine (AP), is not taken. If a reasonably large fraction of the study population, say at least 30%, were untreated with AP until failure or censoring, a quite robust approach is to *(i)* regard a subject as censored at the first time the subject is on AP therapy, *(ii)* redefine Q to be the minimum of time to censoring by competing risks, time to loss-to-follow-up, and time to being on AP therapy, and *(iii)* estimate ψ_0 using the methods of Section 6.2.

If only a small fraction of the study population avoided AP therapy, then one can redefine $A(t)$ to be the joint treatment aerosolized pentamidine and AZT taken at time t and specify a deterministic SNFTM (3.1) that has separate parameters for the AZT effect and for the AP effect as described in Robins et al. (1992, Sec. A2.12). An alternative, and preferred approach, is to specify a direct effect structural nested model as described in Robins and Wasserman (1997). Discussion of these latter models is beyond the scope of this chapter.

7. INFERENCE BASED on INSTANTANEOUS RATE RPSNFTMs

7.1. Difficulties incorporating *a priori* biological knowledge with deterministic SNFTMs

In our simple deterministic SNFTM (3.2), the scientific meaning of the parameter ψ_0 was relatively straightforward; $\exp(-\psi_0)$ was the factor by which continuous treatment extended life. More specifically, Cox and Oakes (1984) suggest interpreting since $\partial T/\partial H(\psi_0) = \exp\{-\psi_0 A(T)\}$ $\exp\{-\psi_0 A(t)\}$ as the relative rate at which real time is being used up compared to baseline time at real time t. Thus, if an individual has U years of baseline time to be used up if treatment is withheld, then the actual time T at which the U years of baseline time will have been used is determined by Equation (3.2). One might hope that the physical interpretation of $\partial T/\partial H(\psi_0)$ as the relative rate at which real time is being used up compared to baseline time would serve in more complex settings to allow us to easily incorporate prior biological knowledge as specific functional form restrictions on our deterministic SNFTM models. However, the following example suggests that this is not the case.

Suppose, based on *a priori* biological understanding, it is known that any treatment received at time t would have no effect on survival unless the subject is destined to fail within the next five weeks without additional treatment. An example would be a setting in which *(i)* failure is death from an infectious disease, *(ii)* if death occurs, it always occurs within five weeks from the time of initial unrecorded subclinical infection, and *(iii)* $A(t)$ is a preventive antibiotic treatment at t which is of no benefit unless the study subject is already infected by t. The challenge then is how to incorporate such biological knowledge into the functional form of a deterministic SNFTM (3.1). We will see that it is difficult to succeed at this challenge if we try to incorporate the biological knowledge directly into a deterministic SNFTM. However, incorporating such biological knowledge is straightforward in a more general class of causal models, the instantaneous-rate (locally) rank preserving structural nested failure time models (instantaneous-rate RPSNFTMs) which model the effect of a final instantaneous blip of treatment on survival. Since each instantaneous-rate RPSNFTM mathematically entails a unique deterministic SNFTM (3.1) as a solution to a particular differential equation, it follows that by solving this differential equation, we can determine the restrictions on the functional form of our deterministic SNFTM (3.1) implied by the restriction that only treatment received within five weeks of failure can affect failure.

7.2. Instantaneous-Rate RPSNFTMs

To describe instantaneous-rate RPSNFTMs, we shall require a number of additional definitions.

Definition: A treatment regime or plan $\overline{a} \equiv a(\bullet) \equiv \{a(t); 0 \leq t < \infty\}$ is a CADLAG (continuous from the right with left hand limits) function on $[0, \infty)$ that is everywhere continuously differentiable except possibly on a countable set of discontinuity points, only finitely many of which are contained in any bounded interval.

Remark: If *(i)* $\overline{A}(T)$ is generated by a marked point process with hazard $\lambda_A\left(t \mid \overline{A}(t^-), \overline{L}(t^-)\right)$, and *(ii)* we define $A(t) \equiv 0$ if $t > T$, then with probability 1, sample paths of the stochastic process $A(t)$ are treatment regimes with discontinuity set the fixed or random jump times for the process depending on whether the process can jump only at fixed discrete times or in continuous time.

We shall assume that $\overline{L}(T)$ as well as $\overline{A}(T)$ have CADLAG sample paths w.p.1. We define the set Dis to be the possibly random set of times at which $\overline{L}(T)$ or $\overline{A}(T)$ are discontinuous.

Definition: Given $\overline{a}$, let $U_{\overline{a}}$ be the (possibly) counterfactual survival time that would be observed if subjects followed treatment regime $\overline{a}$ until failure.

Note that the baseline failure time U is $U_{\overline{a}}$ for the function $\overline{a}$ that is everywhere zero.

Definition: Given $\overline{a} \equiv a(\bullet)$, let $(\overline{a}(t), 0)$ be the regime that agrees with $\overline{a}$ for $u \leq t$ and is zero for $u > t$. $U_{\overline{a}(t),0}$ is the survival time had regime $\overline{a}$ been followed through t and treatment withheld after t.

We assume $U_{\overline{a}}$ obeys the following natural consistency assumptions that essentially assert that the future cannot affect the past.

Consistency Assumption (A): Given $\overline{a}$ and $t > u$, the following are equivalent: $U_{\overline{a}(u),0} > u$, $U_{\overline{a}} > u$, $U_{\overline{a}(t),0} > u$.

Consistency Assumption (B): Given $\overline{a}$ and $\overline{a}^*$ such that $\overline{a}(u) = \overline{a}^*(u)$, $U_{\overline{a}(u),0} = U_{\overline{a}^*(u),0}$.

The following consistency assumption links the counterfactual variables $U_{\overline{a}}$ to the observable variables $\left(T, \overline{A}(T)\right)$.

Consistency Assumption (C):

$$T = U_{\overline{A}(T),0} \text{ w.p.1.} \tag{7.1}$$

The instantaneous-rate RPSNFTM studied in this section requires the assumption of local rank preservation. In the following definition, parts *(i)* and *(ii)* are the substantive parts. Part *(iii)* contains technical assumptions used later.

Definition of Local Rank Preservation: There is local rank preservation w. p. 1 if:
(i) $U_{\overline{a}(t),0}$ is continuous in t w.p.1 and

$$\lim_{\Delta t \downarrow 0} \left\{ U_{\overline{A}(t+\Delta t),0} - U_{\overline{A}(t),0} \right\} / \Delta t = D\left(U_{\overline{A}(t),0}, t \right) \text{ whenever } U_{\overline{A}(t),0} > t \tag{7.2}$$

where

$$D(u,t) \equiv d\left\{ u, t, \overline{L}(t), \overline{A}(t) \right\}$$

(ii) $d\left(u, t, \overline{L}(t), \overline{A}(t) \right) = 0$ if $a(t) = 0$;
(iii) for $t \notin Dis$ on which $\overline{A}(T)$ or $\overline{L}(T)$ is discontinuous, $D(u,t)$ is bounded and its partial derivatives with respect to u and t are bounded and uniformly continuous.

Eq. (7.2) states that if $U_{\overline{A}(t),0} > t$ then, for infinitesimal positive Δt,

$$U_{\overline{A}(t+\Delta t),0} - U_{\overline{A}(t),0} = D\left(U_{\overline{A}(t),0}, t \right) \Delta t \ . \tag{7.3}$$

Now recall that by $\overline{A}(t)$ CADLAG w.p.1, $A(t)$ is constant in $[t, t+\Delta t)$ for Δt sufficiently small. The L.H.S. of (7.3) is the additive increment in survival time attributable to a final blip of treatment $A(t)\Delta t$ in the interval $[t, t+\Delta t)$ administered at dose rate $A(t)$. Thus (7.3) states that the additional increment $D\left(U_{\overline{A}(t),0}, t \right) \Delta t$ is deterministic function of $\overline{L}(t), \overline{A}(t)$ and $U_{\overline{A}(t),0}$. Hence we will refer to $D(u,t)$ as the instantaneous blip function.

To help understand the meaning of the instantaneous blip function $D(u,t)$, consider the infectious disease example of Section 7.1; it follows from (7.3) that the restriction on the instantaneous blip function $D(u,t)$ implied by the biological knowledge that the treatment received $A(t)$ at time t is only harmful or beneficial to those destined to fail by $t+5$ if they receive no further treatment (i.e., to those with $U_{\overline{A}(t),0} - t < 5$) is that

$$D(u,t) = 0 \text{ if } u - t > 5 \ . \tag{7.4}$$

We now define a instantaneous-rate RPSNFT model under the assumption of local rank preservation.

Definition: If the assumption of local rank preservation holds, we say the data follows a instantaneous-rate RPSNFT $D(u,t,\psi)$ if $D(u,t) = D(u,t,\psi_0)$ where ψ_0 is an unknown parameter and $D(u,t,\psi) \equiv d\left(u, t, \overline{L}(t), \overline{A}(t), \psi \right)$ is a known continuously differentiable function of ψ satisfying *(i)* $D(u,t,0) = 0$, *(ii)* $D(u,t,\psi) = 0$ if $A(t) = 0$, and *(iii)* for each fixed value of

ψ, Assumption *(iii)* in the definition of local rank preservation holds.

We will now show that any instantaneous-rate RPSNFTM implies a unique deterministic SNFTM (3.1). We first show that the instantaneous-rate RPSNFTM

$$D\left(u,t,\psi\right)=1-\exp\left\{\psi A\left(t\right)\right\} \tag{7.5}$$

implies the deterministic SNFTM (3.2). Model (7.5) states that the effect of a final instantaneous brief bit of treatment $A\left(t\right)\Delta t$ is to add or subtract $\left[1-\exp\left\{\psi_0 A\left(t\right)\right\}\right]\Delta t$ to a subject's lifetime, so that $\psi_0=0$ implies no effect of treatment on survival. The regularity conditions in Assumption *(iii)* of the definition of Local Rank Preservation and Theorem (2.3) of Section 6 of Loomis and Sternberg (1968) on the existence and uniqueness of solutions to differential equations guarantee that w.p.1 there exists a unique continuous solution $U\left(t\right)\equiv U_{\overline{A}(t),0}$ to the differential equation

$$dU\left(t\right)/dt=D\left\{U\left(t\right),t\right\} \tag{7.6}$$

satisfying consistency assumption C that $U\left(T\right)\equiv T$.

We now solve the differential equations (7.6) corresponding to model (7.5). Integrating $dU\left(t\right)/dt=1-\exp\left\{\psi_0 A\left(t\right)\right\}$, we obtain $U\left(t\right)=t-\int_0^t\exp\left\{\psi_0 A\left(u\right)\right\}du+c$. Imposing the initial conditions $U\left(T\right)=T$ of consistency assumption C, we obtain $c=\int_0^T\exp\left\{\psi_0 A\left(u\right)\right\}du$ so $U\left(t\right)=t+\int_t^T\exp\left\{\psi_0 A\left(u\right)\right\}du$. Hence $U\equiv U\left(0\right)=\int_0^T\exp\left\{\psi_0 A\left(u\right)\right\}du$ reproducing model (3.1) - (3.2) as promised. It is interesting to note that the additive effect $\left\{1-\exp\left[\psi_0 A\left(t\right)\right]\right\}\Delta t$ of the treatment $A\left(t\right)\Delta t$ implies, by model (3.2), a multiplicative effect of constant unit treatment, i.e., $U_{\overline{a}\equiv 1}=\exp\left(-\psi_0\right)U$ where $\overline{a}\equiv 1$ is the regime that always gives unit treatment. The reason for this is a "compound interest effect" of continuous treatment: any additional increment of survival time due to treatment received at t is itself later subjected to treatment, adding a further increment to survival time, etc. Summing the resulting "infinite series" produces the multiplicative effect on survival time of constant treatment.

More generally, for any instantaneous-rate RPSNFTM $D\left(u,t,\psi\right)$, there exists a unique solution $H\left(t,\psi\right)$ to the differential equation

$$\partial H\left(t,\psi\right)/\partial t=D\left(H\left(t,\psi\right),t,\psi\right) \tag{7.7}$$

satisfying the initial condition $H\left(T,\psi\right)=T$. $H\left(t,\psi\right)$ is a function $h\left(t,T,\overline{A}\left(T\right),\overline{L}\left(T\right),\psi\right)$ of ψ and the data. If the RPSNFTM is correctly specified with true value ψ_0, we have by the uniqueness of the solutions to (7.6) and (7.7) that $H(t, \psi_0) = U(t) \equiv U_{\overline{A}(t),0}$. In particular, abbreviating $H(0, \psi_0)$ to $H(\psi_0)$ and $h\left(0, T, \overline{A}(T), \overline{\mathbf{L}}(T), \psi_0\right)$ to $h\left(T, \overline{A}(T), \overline{L}(T), \psi_0\right)$, we obtain the unique deterministic SNFTM $U = H(\psi_0)$ of Eq. (3.1).

It follows that a CAN g-estimator $\widehat{\psi}$ of the parameter ψ_0 of the instantaneous-rate RPSNFTM can be obtained by g-estimation under the assumptions described in Section 2-5 (including the assumption 2.1 of no unmeasured confounders).

Consider next the instantaneous-rate RPSNFTM

$$D(u, t, \psi) = I(u - t < 5)\{1 - \exp\{\psi A(t)\}\} \tag{7.8}$$

which satisfies the assumption (7.4) that treatment at t only affects those destined to fail by $t + 5$ in the absence of further treatment. Integrating (7.6) with $D(u, t) = D(u, t, \psi_0)$ and imposing the initial condition $U(T) = T$, we obtain $U(t) = t + \int_t^T \exp\{\psi_0 A(u)\} du$ for $U(t) - t \leq 5$. For $U(t) - t > 5$, $U(t)$ solves $U(t) = \{U(t) - 5\} + \int_{U(t)-5}^T \exp\{\psi_0 A(u)\} du$, i.e., $\int_{U(t)-5}^T \exp\{\psi_0 A(u)\} du = 5$. It follows that $U \equiv U(0) = \int_0^T \exp\{\psi_0 A(u)\} du$ for $U < 5$ and U satisfies $\int_{U-5}^T \exp\{\psi_0 A(u)\} du = 5$ when $U > 5$. This implies that $\partial U / \partial T = \exp\{\psi_0 A(T)\}$ when $U < 5$ and $\partial U / \partial T = \exp[\psi_0 \{A(T) - A(U - 5)\}]$ when $U \geq 5$. It follows that when $A(u)$ varies with time u, we do not have a closed form expression for U or for $\partial U / \partial T$. Hence, for the corresponding deterministic SNFTM (3.1), we will not have a closed form expression for the function $H(\psi) = h\left(T, \overline{A}(T), \overline{L}(T), \psi\right)$ or its derivative $\partial H(\psi) / \partial T$, although $H(\psi)$ is easily evaluated by numerical means. This "non-obvious" form of $\partial H(\psi_0) / \partial T = \partial U / \partial T$ justifies the remarks of Section 7.1.

8. INSTANTANEOUS-RATE SNFTMs

8.1. Biological implausibility of local rank preservation

Consider two subjects, say i and j, who have identical survival times and covariate and treatment histories $\left(T, \overline{A}(T), \overline{L}(t)\right)$. It follows from the uniqueness of the solution to the differential equation (7.6) that, under the assumption of local rank preservation, the two subjects would have identical survival times U if treatment had been withheld. This assumption is biologically implausible. To see why, again consider the infectious disease example of Section 7.1. Suppose $A(t)$ is the dose of treatment taken at t, treatment has a beneficial biological affect, subject i and j are both infected at time t subject i fails to absorb his/her dose due to gastrointestinal difficulties, while subject j successfully absorbs his/her dose. Then we would expect $U_i = T_i = T_j > U_j$ since subject j but not subject i experiences the benefit of treatment. Dependence of the magnitude of the treatment affect on unmeasured factors such as bioabsorption and genetic endowment is the rule. In this section, we describe the general class of instantaneous-rate SNFT models which allow the magnitude of the treatment affect to depend on unmeasured factors. Specifically, the class of instantaneous-rate SNFTMs, *(i)* does not require that U be a deterministic function of $\left\{T, \overline{A}(T), \overline{L}(T)\right\}$, and *(ii)* contains the instantaneous-rate RPSNFTMs as a subclass. Furthermore, the parameter ψ of a instantaneous-rate SNFTM can be consistently estimated using the g-estimation procedures described previously.

We now define a new function that will allow us to relax the assumption of local rank preservation. Given continuously distributed failure time variates T_1 and T_2 with survivor functions $S_1(u)$ and $S_2(u)$, recall that the quantile-quantile function $v(u) = S_1^{-1}\{S_2(u)\}$ is the unique function $v(u)$ such that $v(T_2)$ has the same distribution $S_1(u)$ as T_1. We now let $U_{\overline{A}(t),0}$ and $U_{\overline{A}(t+h),0}$ play the roles of T_1 and T_2 where, by convention, $A(u) \equiv 0$ if $u > T$. Specifically, let $\mathcal{V}(u,t,h) \equiv v\left(u,t,h,\overline{L}(t), \overline{A}(t)\right)$ be the unique function such that $U_{\overline{A}(t+h),0}$ and $\mathcal{V}\left(U_{A(t),0}, t, h\right)$ have the same conditional distribution given $\overline{L}(t), \overline{A}(t), T > t$. That is,

$$pr\left[U_{\overline{A}(t+h),0} > \mathcal{V}(u,t,h) \mid \overline{L}(t), \overline{A}(t), T > t\right] = pr\left[U_{\overline{A}(t),0} > u \mid \overline{L}(t), \overline{A}(t), T > t\right] .$$

Note $\mathcal{V}(u,t,0) = u$. We now make a smoothness (differentiability) assumption.

Assumption (*): We assume that *(a)* $D(u,t) \equiv \lim_{h \downarrow 0} \{\mathcal{V}(u,t,h) - \mathcal{V}(u,t,0)\}/h$ exists and is bounded for all (u,t) w.p.1 where the function $D(u,t) \equiv d\left(u,t,\overline{L}(t), \overline{A}(t)\right)$ satisfies assumption *(iii)* in the definition of local rank preservation, and *(b)* for $t > x$,

$$pr\left[U_{\overline{A}(t),0} > u \mid \overline{L}(x), \overline{A}(x)\right] \text{ is continuous in } t . \tag{8.1}$$

We have reused the notation $D(u,t)$ in assumption (*) because, under local rank preservation, $D(u,t)$ as just defined is equal to $D(u,t)$ as defined previously. Even without local rank preservation under Assumption (*), we can regard $D(u,t)\Delta t$ as the effect of a last blip of observed treatment $A(t)$ at t sustained for an instantaneous time Δt on quantiles of $U_{\overline{A}(t),0}$. That is, for infinitesimal positive Δt, if, conditional on $\overline{L}(t), \overline{A}(t)$, u is, say, the z^{th} quantile of $U_{\overline{A}(t),0}$, then $u + D(u,t)\Delta t$ is the z^{th} quantile of $U_{\overline{A}(t+\Delta t),0}$. As before, $D(u,t)$ may be discontinuous for $t \in Dis$.

Remark: It is important to note that we no longer assume $U_{\overline{A}(t),0}$ is continuous in t. This is scientifically important since $U_{\overline{A}(t),0}$ will be discontinuous at t if $A(\bullet)$ is exposure to cigarette smoke and a single molecule of benzpyrene inhaled at time t initiates lung cancer. However, Eq. (8.1) remains reasonable since the probability of lung cancer being initiated in $[t, t + \Delta t)$ is small. However, Eq. (8.1) and thus Assumption (*) would be inappropriate if $A(t)$ recorded whether a subject received a mammogram or any other truly "point-source" exposure at time t where, $A(t)$ is a point-source if exposure of $\{t; A(t) \neq 0\}$ is a finite set w.p.1. Models for the effect of point-source exposures, such as mammography, will not be further discussed in this chapter; the SNFTMs discussed in Robins et al. (1992, Appendix 2) are appropriate.

It then follows from Theorem (2.3) of Sec. 6 of Loomis and Sternberg (1968) that *(i)* there exists a unique continuous solution $\mathcal{H}(t) \equiv h\left(t, T, \overline{L}(T), \overline{A}(T)\right)$ to the differential equation $d\mathcal{H}(t)/dt = D(\mathcal{H}(t), t)$ satisfying $\mathcal{H}(T) = T$.

Under local rank preservation, we have seen that this unique solution $\mathcal{H}(t)$ is precisely $U(t) \equiv U_{\overline{A}(t),0}$. This will not be true in the absence of local rank preservation, since $U(t)$ will no longer satisfy (7.6). However, our main result is the following theorem which states that $\mathcal{H}(t)$ and $U(t)$ continue to have the same conditional distributions.

Theorem 8.1: $\mathcal{H}(t)$ and $U_{\overline{A}(t^-),0}$ have the same conditional distribution given $\left(\overline{L}(t), \overline{A}(t), T > t\right)$. In particular, $\mathcal{H} \equiv \mathcal{H}(0)$ and U have the same marginal distributions.

As yet, Theorem (8.1) has only been proved in the special case where the jump times of the $\overline{A}(t)$ and $\overline{L}(t)$ processes are fixed rather than random (Robins, 1998), although it is almost certain that Theorem (8.1) holds in general. The limitation to non-random jump times for the measured L and A processes is no limitation in practice, since we can suppose them to have been measured, say, every second rather than continuously, and then the theorem is true.

We say the data follow a instantaneous-rate SNDM if there is a function $D(u, t, \psi)$ satisfying $D(u, t) = D(u, t, \psi_0)$ with $D(u, t, \psi)$ satisfying the conditions previously described under the definition of a instantaneous-rate RPSNFTMs.

Again letting $H(t, \psi) \equiv h\left(t, T, \overline{A}(T), \overline{L}(T), \psi\right)$ be the solution to the differential equation (7.7) and setting $H(\psi) \equiv H(0, \psi)$, it immediately follows by uniqueness that $\mathcal{H} = H(\psi_0)$. Hence, a CAN estimator $\widehat{\psi}$ of the causal parameter ψ_0 of a instantaneous-rate SNFTM can be obtained by g-estimation under the assumptions described in Sections 2-5.

9. ESTIMATING the DISTRIBUTION of $U_{\overline{a}}$

Given a instantaneous-rate SNFTM $D(u,t,\psi)$, we have shown how to obtain a CAN estimator $\widehat{\psi}$ of ψ_0 under the assumptions of Sections 2-5. However, we often wish to estimate the survival curves $S_{U_{\overline{a}}}(t)$ of $U_{\overline{a}}$ for various treatment regimes $\overline{a}$. Suppose censoring is absent. Then $\widehat{S}_U(t) = n^{-1}\sum_i I\left\{H_i\left(\widehat{\psi}\right) > t\right\}$ is a CAN estimator of $S_U(t)$. The main tool we shall use to estimate other $S_{U_{\overline{a}}}(t)$ is the blip-up function $B(u,t) \equiv b\left(u,t,\overline{L}(t),\overline{A}(t)\right)$ defined to be the unique continuous solution to $dB(t)/dt = D(B(t),t)$ through $(0,u)$.

Example: For model (7.5) with $D(u,t) = 1 - \exp\{\Psi_0 A(t)\}$, we obtain upon integrating that $B(t) = t - \int_0^t \exp\{\psi_0 A(u)\}\,du + c$. By the initial condition $B(0) = u$, we obtain $c = u$, so

$$B(u,t) \equiv B(t) = u + t - \int_0^t \exp\{\psi_0 A(u)\}\,du \ . \tag{9.1}$$

In general, $B(u,t)$ is related to the blip-down function $\mathcal{H}(t)$ by $B(\mathcal{H},t) = \mathcal{H}(t)$ where $\mathcal{H} \equiv \mathcal{H}(0)$. For any covariate and treatment histories $\overline{\ell}$ and $\overline{a}$ defined on $[0,\infty)$, define $b^*\left(u,\overline{\ell},\overline{a}\right)$ to be the solution t^* to $t^* = b\left(u,t^*,\overline{\ell}(t^*),\overline{a}(t^*)\right)$ if one exists and $b^*\left(u,\overline{\ell},\overline{a}\right) \equiv \infty$ otherwise. Note $b^*\left(\mathcal{H},\overline{L},\overline{A}\right) = T$, since $T = B(\mathcal{H},T)$ where $L(u) \equiv A(u) \equiv 0$ if $u > T$.

Example: With $B(u,t)$ given by (9.1), $b^*\left(u,\overline{\ell},\overline{a}\right) = b^*(u,\overline{a})$ is the unique solution to $u = \int_0^{b^*(u,\overline{a})} \exp\{\psi_0 a(u)\}\,du$.

If *(i),* as in model (7.5), $d\left(u,t,\overline{\ell}(t),\overline{a}(t)\right) \equiv d(u,t,\overline{a}(t))$ does not depend on $\overline{\ell}(t)$ [i.e., there is no treatment-covariate interaction] so $b^*\left(u,\overline{\ell},\overline{a}\right) = b^*(u,\overline{a})$, and *(ii)* there are no unmeasured confounders for each $U_{\overline{a}}$, i.e.,

$$U_{\overline{a}} \amalg A(t) \mid \overline{L}(t^-), \overline{A}(t^-), T > t \ . \tag{9.2}$$

Then

$$S_{U_{\overline{a}}}(t) = pr\{b^*(U,\overline{a}) > t\} \ . \tag{9.3}$$

By Theorem 8.1, U can be replaced by $\mathcal{H}$ in (9.3). Robins (1993, Appendix 1) discusses conditions weaker than (9.2) which imply (9.3). It now follows that given a CAN g-estimator $\widehat{\psi}$ of ψ_0, $n^{-1}\sum_i I\left\{b^*\left(H_i\left(\widehat{\psi}\right),\overline{a},\widehat{\psi}\right) > t\right\}$ is a CAN estimator of $S_{U_{\overline{a}}}(t)$ where $b^*(u,\overline{a},\psi)$ is $b^*(u,\overline{a})$ under $\psi = \psi_0$.

Models for Cure:

The fact that $b^*(u,\overline{a})$ can be infinite reflects the possibility of "cure." As an example, suppose U represents the time from diagnosis to death from pancreatic cancer in the absence of treatment. Suppose untreated pancreatic cancer is uniformly fatal so that U is finite w.p.1. Suppose, however, that $pr\{b^*(U,\overline{a})=\infty\}=p\neq 0$. Then a fraction p of the population will be cured under treatment regime $\overline{a}$.

As an example, consider the multiplicative blip model

$$D(u,t,\psi)=(u-t)\,\psi A(t)\ . \tag{9.4}$$

Then, by the formula for solutions to linear first-order differential equations (Loomis and Sternberg, Chapter 6)

$$b(u,t,\overline{a}(t))=\exp\left[\int_0^t \psi_0 a(x)\,dx\right]\left\{u-\int_0^t x\psi_0 a(x)\exp\left[-\int_0^x \psi_0 a(v)\,dv\right]dx\right\}$$

which simplifies, when $\overline{a}$ is the constant dose regime a^*, to $\exp(\psi_0 a^* t)\left[u-(\psi_0 a^*)^{-1}\right]+t+(\psi_0 a^*)^{-1}$. Hence, $b^*(u,\overline{a})=\ell n\left\{-(\psi_0 a^*)^{-1}/\left[u-(\psi_0 a^*)^{-1}\right]\right\}/\{\psi_0 a^*\}$ if $u<(\psi_0 a^*)^{-1}$ and $b^*(u,\overline{a})=\infty$ if $u>(\psi_0 a^*)^{-1}$. Thus, the probability of cure under $\overline{a}$ is the probability that U exceeds $1/\{\psi_0 a^*\}$. The intuition behind this result is that, according to model (9.4), $d(u,t,\overline{a}(t),\psi_0)$ exceeds 1 at $t=0$ if and only if $u\psi_0 a^*>1$. If $d(u,t,\overline{a}(t),\psi_0)>1$, then a blip of treatment a^* at t sustained for duration Δt adds more than Δt years to a subject's survival time. If, as in our model when $u>(\psi_0 a^*)^{-1}$, the instantaneous blip function exceeds 1 for each time t, then the subject is cured.

If the failure time variable is death from all causes, we would want $b^*\left(u,\overline{\ell},\overline{a}\right)$ to be finite for all u, $\overline{\ell}$, and $\overline{a}$ which is guaranteed by having $d\left(u,t,\overline{\ell}(t),\overline{a}(t)\right)<1-\sigma$, $\sigma>0$ for all $u,t,\overline{a}(t),\overline{\ell}(t)$. A natural parameterization of a instantaneous-rate SNFTM that essentially accomplishes this is $D(u,t;\psi)=1-\exp\left\{r\left(u,t,\overline{L}(t),\overline{A}(t),\psi\right)\right\}$ for some function $r(\bullet)$ as in models (7.5) and (7.8).

Covariate-Treatment Interaction:

If $d\left(u,t,\overline{\ell}(t),\overline{a}(t)\right)$ depends on $\overline{\ell}(t)$, we can obtain independent draws from the distribution of $U_{\overline{a}}$ under Assumption (9.2) when the covariate process $L(t)$ only jumps at non-random times, say $0,1,2,\ldots$ as follows.

Step 1: Draw U from its marginal distribution.

Step 2: Draw $L(0)$ from $f\{\ell(0)\mid U=u\}$.

Step 3: Set $m=1$.

Step 4: If $b\left[U, m-1, \overline{L}(m-1), \overline{a}(m-1)\right] \leq m$, set $U_{\overline{a}}$ to $b^*\left(U, \overline{L}, \overline{a}\right)$ where $L(t) = 0$ for $t > m$ and agrees with the drawn $\overline{L}(m^-)$ up to time m. Otherwise, draw $L(m)$ from $f\left[L(m) \mid \overline{L}(m^-), \overline{a}(m^-), U, T > m\right]$, increment m by 1, and return to the start of this step.

To carry out this algorithm in practice, we first obtain a g-estimate $\widehat{\psi}$ of the parameter ψ_0 of an instantaneous-rate SNFTM; draw U from the empirical distribution $\widehat{S}_U(t) = n^{-1}\sum_i I\left(H_i\left(\widehat{\psi}\right) > t\right)$; replace the functions $b(\bullet)$ and $b^*(\bullet)$ by $b\left(\bullet, \widehat{\psi}\right)$ and $b^*\left(\bullet, \widehat{\psi}\right)$; and estimate the density $f\left[\ell(m) \mid \overline{\ell}(m^-), \overline{a}(m^-), U, T > m\right]$ by specifying a parametric model $f\left[\ell(m) \mid \overline{\ell}(m^-), \overline{a}(m^-), U, T > m; \eta\right]$ and evaluating it at $\widehat{\eta}$ which maximizes

$$\prod_{i=1}^{n}\prod_{m=0}^{int(T_i)} f\left[L_i(m) \mid \overline{L}_i(m^-), \overline{A}_i(m^-), H_i\left(\widehat{\psi}\right), T_i > m; \eta\right].$$

A heuristic explanation of the above algorithm is as follows. If a simulated subject with baseline time U manages to survive to time m under regime $\overline{a}$, we randomly draw $L(m)$ and then use the blip-up function $b\left(u, t, \overline{\ell}(t), \overline{a}(t)\right)$ to determine whether the subject has survived to time $m+1$ or whether the subject has died at a time $U_{\overline{a}}$, determined by the function $b^*\left(u, \overline{\ell}, \overline{a}\right)$ in the interval $(m, m+1]$. This explanation is heuristic in that it implicitly but unnecessarily assumes local rank preservation. Robins et al. (1992, Appendix 2) discusses how to generalize the results of this section to allow for censoring. A drawback of SNFTMs is that in the presence of a covariate-treatment interaction $S_{U_{\overline{a}}}(t)$ cannot be calculated without modelling the law of $\mathbf{L}(m)$ given $\left\{\overline{\mathbf{L}}(m^-), \overline{A}(m^-), U, T > m\right\}$. Estimation of the non-nested structural Cox proportional hazard models for $U_{\overline{a}}$ described in Appendix 2 can obviate this problem.

Appendix 1: Calculation of the variance of the g-test statistic numerator:

We shall require some notation. Given a stochastic process $G(\bullet)$, let $U_A\{G(\bullet)\} = \int_0^X dM_A(u)\left\{G(u) - E\left[e^{\alpha' W(u)}G(u)\right]/E\left[e^{\alpha' W(u)}\right]\right\}$ where $\alpha' W(u)$ is from model (2.4), $dM_A(u) = dN_A(u) - \lambda_0(u)e^{\alpha' W(u)}du$ and $N_A(u)$ counts the number of jumps in the $A(u)$ process through time u.

Now define $U_1 \equiv U_A\{\tau Q(\bullet, \psi_0)/K(X)\}$ and $U_2 \equiv U_A\{\mathbf{W}(\bullet)\}$. Then define $V_1 \equiv E\left[\left\{U_1 - E\left[U_1 U_2'\right]E\left[U_2^{\otimes 2}\right]^{-1}U_2\right\}^{\otimes 2}\right]$. V_1 is the "robust variance" of the g-test numerator (i.e., Cox partial likelihood score test numerator) of the hypothesis that $\theta = 0$ in the extended Cox model (2.4) when the true $K(X)$ is used.

Now define $V \equiv V_1 - V_{corr}$ where V_{corr} is the correction to the variance V_1 required when we replace $K(X)$ by its estimator $\widehat{K}(X)$. Specifically, $V_{corr} = V_2 + V_3$ where $V_2 = E\left[\int_0^\infty dN_Q(u)\left\{\mathcal{L}^Q(u, J(u))\right\}^{\otimes 2}\right]$, $N_Q(u) = I[Q \leq u, \tau = 0]$, $J(u) \equiv \int_0^X dM_A(t)Q(t, \psi_0)/K(u)$ and, for any $G(u)$, $\mathcal{L}^Q\{u, G(u)\} \equiv E\left[K(u)\{K(X)\}^{-1}\tau G(u)I(X^* > u)e^{\alpha^{*\prime}W^*(u)}\right]/$

$E\left[I\left(X^{*}>u\right)e^{\alpha^{*\prime}W^{*}(u)}\right]$ with $\alpha^{*\prime}W^{*}(u)$ from model (6.3).

$V_3=V_{31}\left\{V_{32}\right\}^{-1}V_{31}^{\prime}$, $V_{31}=E\left[\int dN_Q(u)\left\{\mathcal{L}^Q\left\{u,J(u)W^{*}(u)\right\}-\mathcal{L}^Q\left(u,J(u)\right)\mathcal{L}^Q\left(u,W^{*}(u)\right)\right\}\right]$ and V_{32} is the expected partial information matrix for α^{*} from Cox model (6.3). Since V_{corr} is non-negative definite, the robust variance V_1 is greater than or equal to the true variance V in the non-negative definite sense. A consistent estimator $\widehat{V}$ of V is obtained by substitution into the above formulae according to the following six steps.

(i) Replace any expectation by a sample average over the n study subjects.

(ii) Replace α and α^{*} by their partial maximum likelihood estimates.

(iii) Replace ψ_0 by the value of ψ being tested in the g-test.

(iv) Replace $K(\bullet)$ by $\widehat{K}(\bullet)$.

(v) Replace $dM_A(u)$ by $dN_A(u)-d\widehat{\Lambda}_0(u)e^{\widehat{\alpha}^{\prime}W(u)}$ where $\widehat{\Lambda}_0(u)$ is the Cox cumulative hazard estimate from the model (2.4).

(vi) Estimate V_{32} by the observed partial information matrix from the fit of model (6.3).

Now let $\widehat{U}_1(\psi)$ be U_1 with the substitutions described in steps *(i)* - *(vi)* above. The g-statistic numerator is precisely $\sum_{i=1}^{n}\widehat{U}_{1i}(\psi)$. We then have the following theorem on which our inferences are based.

Theorem: Given that Assumptions (2.2) and (6.2) and models (2.4), (3.1), and (6.3) are correct, then, when $\psi=\psi_0$, $n^{-\frac{1}{2}}\sum_i\widehat{U}_{1i}(\psi)$ is asymptotically normal with mean zero and asymptotic variance V that can be consistently estimated by $\widehat{V}$.

The proof is analogous to that given on page 284-285 of Robins (1993) using the methods developed in Robins and Rotnitzky (1992).

Appendix 2: A (non-nested) structural Cox proportional hazard model specifies

$$\lambda_{U_{\overline{a}}}(t)=\lambda(t)\exp\left[r\left\{t,\overline{a}\left(t^{-}\right),\beta_0\right\}\right] \tag{33}$$

where $r(\cdot)$ is a known function and $\lambda(t)$ an unspecified baseline hazard. For simplicity, assume $A(t)$ is dichotomous. Then a consistent asymptotically normal estimator $\widehat{\beta}$ of β_0 under assumptions (18), (30) and models (4), (19), and (33) is the solution to the weighted Cox score equation for T

$$0=\sum_i\int_0^{\infty}\left\{dN_{Ti}(u)/\widehat{\Omega}_i(u)\right\}\left\{P_i(u,\beta)-\widehat{E}\left[P(u,\beta),\beta\right]/\widehat{E}\left[\mathbf{1}(u),\beta\right]\right\}$$

where *(i)* $N_T(u)=I\left[X^{*}\leq u,X^{*}=T\right]$, *(ii)* $P(u,\beta)=p\left[u,\overline{A}\left(u^{-}\right),\beta\right]$ is a vector function of $\dim\beta$ chosen by the analyst such as $\partial r\left(u,\overline{A}\left(u^{-}\right),\beta\right)/\partial\beta$, *(iii)* $\mathbf{1}(u)=1$ is the identity, *(iv)* $\widehat{E}\left[J(u,\beta),\beta\right] = \sum_i I\left(X_i^* > u\right)\exp\left[r\left\{u,\overline{A}\left(u^-\right),\beta\right\}\right]J_i(u,\beta)/\widehat{\Omega}_i(u)$, *(v)* $\widehat{\Omega}_i(u) = \widehat{K}_i(u)\widehat{K}_{Ai}(u)$, $\widehat{K}(u)$ as defined in the text, (vi) $\widehat{K}_A(u) =$

$\exp\left[-\int_0^u \widehat{\lambda}_A\left[t \mid \overline{L}\left(t^-\right),\overline{A}\left(t^-\right)\right]dt\right]\prod_{\{t;t<\mu \text{ and } A(t)\neq A(t^-)\}}\widehat{\lambda}_A\left(t \mid \overline{L}\left(t^-\right),\overline{A}\left(t^-\right)\right)$ where

$\widehat{\lambda}_A\left[t \mid \overline{L}\left(t^-\right),\overline{A}\left(t^-\right)\right] = \widehat{\lambda}_0(u)\exp\left[\widehat{\alpha}'\mathbf{W}(u)\right]$ and $\widehat{\lambda}_0(u)$ is now a kernel smoothed version of the Cox estimate of $\lambda_0(t)$ of model 4 as in Ref.26. $\widehat{\beta}$ combined with the estimate $\widehat{\lambda}(t) = \sum_i dN_{Ti}(t)/\left\{\widehat{\Omega}_i(t)\widehat{E}\left[\mathbf{1}(t),\widehat{\beta}\right]\right\}$ of $\lambda(t)$ produces an estimate of $\lambda_{U_{\overline{a}}}(t)$ and thus of $S_{U_{\overline{a}}}(t)$. In the above, we have assumed that Q is the minimum of time to loss to follow-up, competing risk, and end to follow-up as discussed in the remark in the section "Censoring by Competing Risks." Note $\widehat{K}_A(u)$ is a consistent estimate of probability that a subject would have his observed history $\overline{A}(u)$ through time u.

Bibliography

1. Arjas, E. (1989). Survival models and martingale dynamics (with discussion). *Scandinavian Journal of Statistics*, 15:177-225.

2. Arjas, E. and Eerola, M. (1993). On predictive causality in longitudinal studies. *Journal of Statistical Planning and Inference*, 34:361-384.

3. Dawid, A.P. (1979). Conditional Independence in Statistical Theory (with discussion). *Journal of the Royal Statistical Society*, Ser. B, 41:1-31.

4. Holland P.W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81:945-960.

5. Klein, J.P., Keiding, N., and Copelan, E. (1994). Plotting summary predictions in multistate survival models: Probabilities of relapse in death and remission for bone marrow transplantation patients. *Statistics in Medicine*, 13:2315-2332.

6. Loomis, B., and Sternberg, S. (1968). **Advanced Calculus**. Addision Wesley.

7. Mark S.D., and Robins, J.M. (1993a). Estimating the causal effect of smoking cessation in the presence of confounding factors using a rank preserving structural failure time model. *Statistics in Medicine*, 12:1605-1628.

8. Mark S.D., and Robins, J.M. (1993b). A method for the analysis of randomized trials with compliance information: An application to the multiple risk factor intervention trial. In: **Controlled Clinical Trials.** 14:79-97. New York: Elsevier.

9. Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82:669-690.

10. Robins, J.M. (1986). A New Approach to Causal Inference in Mortality Studies with Sustained Exposure Periods – Application to Control of the Healthy Worker Survivor Effect. *Mathematical Modelling*, 7:1393-1512.

11. Robins, J.M. (1987). Addendum to "A New Approach to Causal Inference in Mortality Studies with Sustained Exposure Periods – Application to Control of the Healthy Worker Survivor Effect." *Computers and Mathematics with Applications*, 14:923-945.

12. Robins, J.M. (1989). The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In: **Health Service Research Methodology: A Focus on AIDS**. NCHSR, U.S. Public Health Service. Eds: Sechrest, L., Freeman, H., Mulley, A. pp. 113-159.

13. Robins, J.M., Tsiatis, A. (1991). Correcting for non-compliance in randomized trials using rank-preserving structural failure time models. *Communications in Statistics,* 20:2609-2631.

14. Robins, J.M. (1992). Estimation of the time-dependent accelerated failure time model in the presence of confounding factors. *Biometrika,* 79:321-34.

15. Robins, J.M., Blevins, D., Ritter, G, and Wulfsohn, M. (1992). G-estimation of the effect of prophylaxis therapy for pneumocystis carinii pneumonia on the survival of AIDS patients. *Epidemiology* 3:319-336.

16. Robins, J.M., and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In: **AIDS Epidemiology - Methodological Issues**. Eds: Jewell, N., Dietz, K., Farewell, V. Boston, MA: Birkhäuser. pp. 297-331.

17. Robins, J.M. (1993). Analytic methods for estimating HIV treatment and cofactor effects. In: **Methodological Issues of AIDS Mental Health Research**. Eds: Ostrow, D.G., Kessler, R. New York: Plenum Publishing. pp. 213-290.

18. Robins, J.M., and Greenland S. (1994). Adjusting for differential rates of PCP prophylaxis in high- versus low-dose AZT treatment arms in an AIDS randomized trial. *Journal of the American Statistical Association*, 89:737-749.

19. Robins, J.M. (1994). Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics*, 23:2379-2412.

20. Robins, J.M. (1995). Estimating the Causal Effect of a Time-varying Treatment on Survival using Structural Nested Failure Time Models. (To appear, *Statistica Neederlandica*).

21. Robins, J.M. (1997). Causal inference from complex longitudinal data. In: **Latent Variable Modeling and Applications to Causality. Lecture Notes in Statistics (120)**, M. Berkane, Editor. NY: Springer Verlag, 69-117.

22. Robins, J.M. (1997). Estimation and testing of direct effects by reparameterizing directed acyclic graphs. **Causation and Computation**, ed. G. Cooper and C. Glymour, MIT/AAAI Press, 1998 (to appear).

23. Robins, J.M. (1998). Correction for non-compliance in bioequivalence trials. *Statistics in Medicine* (To appear).

24. Van der Laan, M.J. and Hubbard, A. (1997). Estimation with interval-censored data and covariates. *Lifetime Data Analysis* 3:77-91.

25. Rosenbaum, P.R. (1984). Conditional permutation tests and the propensity score in observational studies. *Journal of the American Statistical Association*, 79, 565-574.

26. Rosenbaum, P.R. (1984). The consequences of adjustment for a concomitant variable that has been adversely affected by treatment. *Journal of the Royal Statistical Society A*, 147, 656-666.

27. Rubin, D.B. (1978). Bayesian Inference for Causal Effects: The Role of Randomization. *The Annals of Statistics*, 6:34-58.

28. Spirtes, P., Glymour, C., and Scheines, R. (1993). **Causation, Prediction, and Search**. New York: Springer Verlag.

29. Lin DY, Wei LJ. (1989). The robust inference for the Cox proportional hazard model. *Journal of the American Statistical Association* 84:1074-1078.

30. Witteman, J.C.M., D'Agogostino, R.B., Stijnen, T., Kannel, W.B., Cobb, J.C., de Ridder, M.A.J., Hofman, A., and Robins, J.M. (1996). G-estimation of causal effects: Isolated systolic hypertension and cardiovascular death in the Framingham Study. (Submitted to the *American Journal of Epidemiology.)*