Attenuation caused by infrequently updated covariates in survival analysis

PER KRAGH ANDERSEN*, KNUT LIESTØL†

Department of Biostatistics, University of Copenhagen, Blegdamsvej 3, DK 2200 Copenhagen N, Denmark
pka@biostat.ku.dk

SUMMARY

This paper deals with hazard regression models for survival data with time-dependent covariates consisting of updated quantitative measurements. The main emphasis is on the Cox proportional hazards model but also additive hazard models are discussed. Attenuation of regression coefficients caused by infrequent updating of covariates is evaluated using simulated data mimicking our main example, the CSL1 liver cirrhosis trial. We conclude that the degree of attenuation depends on the type of stochastic process describing the time-dependent covariate and that attenuation may be substantial for an Ornstein–Uhlenbeck process. Also trends in the covariate combined with non-synchronous updating may cause attenuation. Simple methods to adjust for infrequent updating of covariates are proposed and compared to existing techniques using both simulations and the CSL1 data. The comparison shows that while existing, more complicated methods may work well with frequent updating of covariates the simpler techniques may have advantages in larger data sets with infrequent updatings.

Keywords: Attenuation; Cox regression model; Measurement errors; Survival analysis; Time-dependent covariates.

1. INTRODUCTION

In survival analysis, regression models are frequently specified via the hazard function, one advantage being that for some such models it is possible to include covariates that change over time. Time-dependent covariates may be used as a basis for tests of the validity of assumptions of the model (e.g. proportional hazards), for handling treatment switches and for utilizing updated information on time-varying variables recorded during follow-up of individuals.

We here consider updated, quantitative covariates. Ideally, such covariates should be continuously registered with error-free measurements. However, measurement errors will often affect covariate values and it is well known (e.g. Prentice, 1982) that such errors will tend to attenuate regression coefficients. Moreover, markers of disease development are normally only measured at (possibly infrequent and irregularly spaced) follow-up times. Thus, there will be unobserved stochastic changes in the covariate from the time of measurement to the times for which information is needed in the estimation procedure. Such changes that arise due to the ‘ageing’ of measurements may be viewed as a kind of measurement error, suggesting that we should expect attenuation of regression coefficients. However, depending on the kind of stochastic processes that describe a covariate, the statistical properties of the error that arises due

*To whom correspondence should be addressed. Also at: Danish Epidemiology Science Centre, Copenhagen, Denmark.
†Also at: Department of Informatics, University of Oslo, Norway.

Biostatistics 4(4) © Oxford University Press; all rights reserved.
to the ageing will differ markedly, and analogies to different models for measurement errors suggest quite different levels of attenuation.

The appropriateness of adjustment for measurement errors will depend on the purpose of the study. If prediction of the outcome based on available explanatory variables is the issue then one rarely needs to worry about measurement errors (Carroll et al., 1995, p. 19). However, when the purpose is to understand a biological relationship between the explanatory variables and the outcome then adjustment for measurement errors is important. For the effect of ageing the situation is quite different. Unless data adhere to a strict follow-up scheme whereby synchronous updating of covariates for all individuals is obtained and prediction is the purpose of the study then some sort of adjustment for ageing of covariates is warranted.

A number of approaches to compensate for infrequent updating of covariates have been discussed previously. Thus, one set of methods is based on smoothing the observed updated covariates and thereby predicting the value at any time (Raboud et al., 1993; Tsiatis et al., 1995; Boscardin et al., 1998; Bycott and Taylor, 1998; Dafni and Tsiatis, 1998). Another set of methods uses joint modelling of the time-dependent covariate and survival (DeGruttola and Tu, 1994; Faucett and Thomas, 1996; Wulfsohn and Tsiatis, 1997; Henderson et al., 2000; Xu and Zeger, 2001). Recently, Tsiatis and Davidian (2001) suggested a conditional score procedure for estimating the parameters in the hazard model.

We shall both discuss why attenuation may arise and suggest some alternative, simple methods for estimation of the effect of the updated covariate. These techniques, though motivated by a model-based discussion, are quite intuitively based and no rigorous analysis of their properties is intended. Rather, we will compare these simple techniques with some of the above-mentioned approaches using simulated data and a real-life example, the CSL1 trial in liver cirrhosis. The structure of the paper is as follows. Section 2 describes the CSL1 trial while in Section 3.1 we first review models with time-dependent covariates, the main example being the Cox (1972) proportional hazards model. We then discuss models for the development of the time-dependent covariate. Section 3.3 describes the simulations. A model-based discussion of why attenuation may occur is given in Section 4, methods to adjust for attenuation are discussed in Section 5 and illustrated using both simulated data and CSL1. Some final comments are given in Section 6.

2. THE CSL1 LIVER CIRRHOSIS TRIAL

CSL1 was a double blind multicentre randomized clinical trial with the purpose of studying the effect of prednisone treatment versus placebo on survival with liver cirrhosis (CSL, 1974). The accrual period ranged from 1962 to 1969 and the patients were followed to death, censoring or to September 1974. Follow-up examinations were scheduled at entry, after 3, 6 and 12 months and thereafter once a year. The actual follow-up times showed some deviations from this schedule, and we here use a cleaned data set where measurements that could be identified as taken at unscheduled follow-up visits are removed (Liestøl and Andersen, 2002). The variables recorded at follow-up included the prothrombin index based on measurements of coagulation factors II, VII and X produced by the liver and given as ‘per cent of normal’, a low value indicating impaired liver function. The analyses by Schlichting et al. (1983) and Christensen et al. (1985), using the Cox regression model with time-fixed variables, included the same 488 patients used in this study. The number of deaths was 292. Christensen et al. (1986) analysed the data using time-dependent variables. Further analyses studying the prothrombin index were presented by Andersen et al. (1993, Section VII.2) and Andersen (1986).

The CSL1 data will be used both as illustration (Section 5) and as a basis for simulations (Section 3.3). For ease of presentation, we focus on models including only the prothrombin index and age at entry (in years). The estimated effects of these variables are only moderately changed by adding other prognostic
factors, including the treatment indicator which turned out to be insignificant whether or not adjustment for prothrombin was made.

With these two variables included as time-fixed covariates the estimates are (with standard errors in brackets) $\hat{\beta}_{\text{pro}} = -0.015(0.0027)$, $\hat{\beta}_{\text{age}} = 0.046(0.0065)$. A model with prothrombin index as a time-dependent covariate (using for robustness a lagged value: at time $t$ the last value recorded prior to time $t - 7$ days) gives $\hat{\beta}_{\text{pro}} = -0.025(0.0025)$, $\hat{\beta}_{\text{age}} = 0.049(0.0067)$. The interpretation and ‘correctness’ of these estimates will be discussed in Section 5 where further analyses of the CSL1 data are presented.

3. HAZARD REGRESSION MODELS WITH TIME-DEPENDENT COVARIATES

3.1 Hazard models

Assume that independent individuals have life times $T_i^*$, $i = 1, \ldots, n$ whose distribution is assumed to depend on vectors $(Z_i^*(t), t \geq 0)$ of, possibly time-dependent, covariates. Frequently, practical restrictions prevent the observation of all $T_i^*$ and, furthermore, all individuals may not be observed from the same time origin. In this case the data will consist of entry times $V_i \geq 0$, $i = 1, \ldots, n$, possibly right-censored times $T_i$, indicators $D_i, \ldots, D_n$ of failure, and the covariate histories $(\{Z_i^*(u), V_i \leq u \leq T_i\}, i = 1, \ldots, n)$ from the time of entry to the time of exit.

The relation between the covariates and survival is specified via the hazard function for $T_i^*$, i.e.

$$\alpha_i(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}(T_i^* \leq t + \Delta t \mid T_i^* \geq t, \{Z_i^*(u), 0 \leq u \leq t\})}{\Delta t}. $$

The main example of such a hazard model is the Cox regression model

$$\alpha_i(t) = \alpha_0(t) \exp(\beta_i Z_i(t)).$$

(3.1)

Here, $Z_i(t)$ is a numerically coded function of the covariate history $\{Z_i^*(u), u \leq t\}$ up to time $t$. Thus, $Z_i(t)$ may simply be taken as the present value $Z_i^*(t)$, as a suitably ‘lagged’ value $Z_i^*(t - d)$ or as a more complicated function of the history at time $t$, as in Gail (1981). The baseline hazard $\alpha_0(t)$ is left unspecified.

In this model, the regression coefficients, $\beta_i$ are estimated by the value $\hat{\beta}$ maximizing the Cox partial likelihood

$$L(\beta) = \prod_{i=1}^n \left( \frac{\exp(\beta_i Z_i(T_i))}{\sum_{j \in R(T_i)} \exp(\beta_j Z_j(T_i))} \right)^{D_i}.$$ 

(3.2)

and the cumulative baseline hazard $\hat{A}_0(t) = \int_0^t \alpha_0(u) du$ by the Breslow estimator

$$\hat{A}_0(t) = \sum_{T_i \leq t} \frac{D_i}{\sum_{j \in R(T_i)} \exp(\hat{\beta}_j Z_j(T_i))}.$$ 

(3.3)

In (3.2)–(3.3), $R(T_i)$ is the set of all individuals under observation just before time $T_i$, i.e. the risk set $\{j : V_j < T_i \leq T_j\}$. We will assume throughout that the mechanisms of right-censoring and delayed entry are independent (Andersen et al., 1993, Chapter III) whereby consistent estimation of the model parameters is possible. Note, however, that in order to compute (3.2)–(3.3) the covariate values $Z_j(T_i)$ at the failure time $T_i$ are needed for all individuals $j \in R(T_i)$ whereas the covariate processes $Z_j^*(t)$ are typically only recorded at follow-up visits and not continuously. This is why the ageing problem occurs for this model leading, possibly, to inconsistent estimators.
An alternative hazard model where time-dependent covariates may be included is the non-parametric additive model of Aalen (1980, 1989) given by

$$\alpha_i(t) = \alpha_0(t) + \beta(t)Z_i(t),$$  
(3.4)

where both the baseline hazard $\alpha_0(t)$ and the regression functions $\beta(t)$ are left unspecified. A semi-parametric version of (3.4) with time-independent regression functions, $\beta_j(t) = \beta_j$, and unspecified baseline hazard was discussed by Lin and Ying (1994):

$$\alpha_i(t) = \alpha_0(t) + \beta Z_i(t).$$  
(3.5)

Also in the models (3.4)–(3.5), observation of the time-dependent covariates is needed at all event times in order to estimate the regression parameters consistently.

In what follows we shall restrict attention to a single time-dependent covariate, $Z_i(t)$. Extensions to multivariate covariate processes are possible but we shall omit the details. Assume that, in addition to $Z_i(t)$, the model contains time-fixed and error-free covariates, $x_i$, that is, like the model studied by Tsiatis and Davidian (2001), our Cox model (3.1) becomes

$$\alpha_i(t) = \alpha_0(t) \exp(\beta Z_i(t) + \delta'x_i).$$  
(3.6)

The additive models (3.4) and (3.5) may be modified similarly. Following Section 2 we shall refer to $Z_i(t)$ and $x_i$ as ‘prothrombin’ and ‘age’, respectively.

### 3.2 Covariate model

We use a covariate model that appears reasonable as a rough model for prothrombin in CSL1 and in addition is suitable when discussing causes of attenuation. Similar models have been used in other applications, e.g. for CD4 counts in HIV-positive individuals (Taylor and Law, 1998). Thus, the true coded prothrombin value $Z_i(t)$ for individual $i$ at time $t$ is described as

$$Z_i(t) = \mu(t) + A_i + U_i(t) \tag{3.7}$$

where the components are independent and have the following interpretation:

- $\mu(t)$: Common deterministic trend. A common trend will not affect parameter estimates in the Cox model if updating is continuous or synchronous since a factor of the form $e^{\beta \mu(t)}$ will cancel out in the partial likelihood (3.2). With non-synchronous updating, this will no longer be true.
- $A_i$: Individual level; $A_i \sim N(0, \omega^2)$ for patient $i$.
- $U_i(t)$: A stochastic process. Two alternatives are studied. A Brownian motion with no systematic drift is well suited for illustrative purposes due to its simplicity. It has $E(U_i(t + \Delta)|U_i(t)) = U_i(t)$, a variance proportional to $t$, and correlation of the form $\text{corr}(U_i(t), U_i(t + \Delta)) = \rho(t, \Delta) = \sqrt{\Delta/(t + \Delta)}$. However, it has clear limitations as a biological model since the variability of biochemical substances will typically be limited by homeostatic mechanisms, e.g. prothrombin in CSL1 shows a fairly stable variance (Liestøl and Andersen, 2002). To simulate this, we may use stationary or near-stationary processes. Thus, a possible model is the so-called Ornstein–Uhlenbeck (OU) process, which is stationary with a variance $\sigma^2$ and a correlation of the form $\rho(t, \Delta) = \exp(-\gamma \Delta)$. A more complex process like an integrated OU-process (Taylor and Law, 1998) might refine the approximation to the CSL1 data, but the simpler OU-process is more suitable for illustrative purposes and still not too far from reality.
Furthermore, errors will arise due to sub-optimal technical procedures and most biochemical markers show short-term variations of minor prognostic value. We denote the combined effect of such factors as measurement errors, \( \varepsilon_{i}(t) \). We assume \( \varepsilon_{i}(t) \sim N(0, \tau^2) \), that all errors are independent of each other and of all other model components, and that the effect is additive. Thus, for individual \( i \) at time \( t \), the observed value \( W_{i}(t) \) is given by

\[
W_{i}(t) = Z_{i}(t) + \varepsilon_{i}(t)
\]

and the data for prothrombin will then be \( W_{i}(t_{ij}) \) where \( 0 \leq V_{i} = t_{0} < t_{1} < \cdots < t_{n_{i}} \leq T_{i} \).

### 3.3 Simulation model

In simulations we use ‘caricatures’ of the CSL1 study. As the true hazard model, we use a discrete time approximation to the Cox model (3.6). The estimated baseline intensity for CSL1 is fairly constant (e.g. Andersen, 1986). However, procedures for handling updated covariates may behave well with constant intensity, yet be unacceptable under alternative intensities. Our standard simulations with constant intensity (\( \alpha(t) = 0.1 \)) are therefore supplemented using baseline intensities of the Gompertz–Makeham form \( \alpha(t) = p_{0} + p_{1} \exp(p_{2}t) \), including a sharply increasing intensity obtained with \( p_{0} = -0.05, p_{1} = 0.05 \) and \( p_{2} = 0.25 \). Censoring is independent of the failure time process and occurs with constant intensity adjusted to obtain about 40% censoring, roughly as in CSL1.

Similarly, the Brownian motion is approximated by a random walk with Normally distributed steps. The OU-process is approximated by the difference equation \( U_{i}(k) = (1 - \gamma)U_{i}(k-1) + V_{i}(k) \), where \( U_{i}(k) \) is the value at the \( k \)th iteration step for individual \( i \), and \( V_{i}(k) \sim N(0, \sigma_{i}^{2}) \). Moreover, \( U(0) \sim N(0, \sigma_{0}^{2}/(2\gamma - (\gamma)^{2})) \), corresponding to the stationary distribution.

If we assume that models (3.7) and (3.8) are true with a linear function for \( \mu(t) \) and an OU-process for \( U_{i}(t) \), the covariance structure is known and the data may be analysed using a mixed linear model (e.g. PROC MIXED in SAS). The CSL1 data were analysed using this model with an added indicator for treatment. Selective removal of high-risk individuals may bias the estimates but simulations (Section 4) suggest that this is of minor importance. We then obtained variance estimates of \( \hat{\sigma}^{2} = 195, \hat{\tau}^{2} = 274, \hat{\gamma} = 0.282 \) (time scale in years) and \( \hat{\omega} = 72 \) (per cent) with a treatment effect of +13. These figures, except for the non-simulated treatment effect, were used as standard parameter values in our simulations. The same values were used when simulating without measurement errors, except \( \tau = 0 \), and when simulating with a Brownian motion for \( U_{i}(t) \), except that \( \gamma = 0 \) and \( \omega^{2} = 300 \) to compensate for the initial zero variance of the Brownian motion.

Age at entry is assumed independent of prothrombin and Normally distributed with mean 60 (years) and standard deviation 10 (years). Standard sample size is 500 (488 in CSL1), while supplementary analyses used 100, 5000 and 10000 individuals. The coefficient for prothrombin is set to \(-0.04\) and for age to 0.05.

### 4. ATTENUATION OF COVARIATES

Figure 1(a) illustrates the effects of measurement errors and infrequent updating using the simulation model mimicking CSL1. The two curves correspond to models with and without measurement error, while the individual crosses are obtained with different updating frequencies. The figure shows pronounced effects both of measurement errors (the difference between the two curves) and of ageing of covariates (the changes along the abscissa).

In this section we discuss when and why attenuation may occur based on particular models for the covariate process and illustrate it using simulations. In Section 5, motivated by these results, we discuss ways to adjust for attenuation.
Fig. 1. (a) Attenuation due to measurement errors and ageing of covariate measurements. Simulations have been carried out with different updating frequencies and with (o) and without (x) measurement errors. The fitted curves are of the form \(a + be^{-ct}\). The figure is based on a simulation with 5000 individuals and the standard OU-model of Section 3.3. (b) Deletion-extrapolation procedure applied to the CSL1 data. The upper curve is a fitted second order polynomial. The point labelled C is obtained by censoring observations one year after measurement (and reentering the patient after the next measurement). The lower curve is obtained by multiplying the upper curve by the measurement error attenuation factor 0.69 estimated in Section 5. The circles on this curve are SIMEX estimates.

We assume that a hazard model \(\alpha(t \mid Z_i(t), x_i)\) like (3.6), or a similar version of (3.4)–(3.5), holds. Under the further assumption that \(\alpha(t \mid Z_i(t), x_i, W_i(s), s \leq t) = \alpha(t \mid Z_i(t), x_i)\), i.e. the hazard function does not depend on \(W_i(\cdot)\) when \(Z_i(t)\) is known, it can be shown (Andersen et al., 1993, Section II.4.2; Prentice, 1982) that the hazard function given only the covariate measured with error at an earlier time \(t - \Delta\) (dropping the index \(i\)) is

\[
\alpha(t \mid W(t - \Delta), x) = E\left(\alpha(t \mid Z(t), x) \mid W(t - \Delta), x, T^* \geq t\right).
\]
For the Cox model (3.6) this becomes

\[
\alpha(t \mid W(t - \Delta), x) = \alpha_0(t) \exp(\beta W(t - \Delta) + \delta^\prime x) \\
\times E\left(\exp(\beta(Z(t) - Z(t - \Delta))) \mid W(t - \Delta), x, T^* \geq t\right) \\
\times E\left(\exp(-\delta^\prime(t - \Delta)) \mid W(t - \Delta), x, T^* \geq t\right). 
\] (4.1)

If we estimate using the Cox model with error-prone and aged covariate values, we will clearly not estimate \(\beta, \delta\) since the model lacks the two last factors of (4.1). It is well known from studies on frailty models (e.g. Bretagnolle and Huber-Carol, 1988; Keiding et al., 1997) or on measurement errors (e.g. Carroll et al., 1988) that omission of such factors causes attenuation. Note that the last but one factor in (4.1) reflects ageing (equal to 1 if \(\Delta = 0\)), while the last reflects the effect of the measurement errors. For the pure measurement error situation (\(\Delta = 0\)) Aalen (1989) showed that a similar attenuation occurs for the additive model (3.4), while a general discussion of measurement errors in the additive model (3.5) was given by Kulich and Lin (2000).

In order to evaluate (4.1) (or a similar version of the additive hazard model) analytically we need the joint distribution of \(x, Z(t)\) and \(W(t - \Delta)\) among survivors at \(t\), i.e. the conditional distribution given \([T^* \geq t]\). Due to the selective removal over time of high-risk individuals this distribution will typically differ from the marginal distribution of \((x, Z(t), W(t - \Delta))\) and, with the exception of some specific situations (Aalen, 1989; Martinussen and Keiding, 1997), it is difficult to assess. Here, we shall follow Tsiatis et al. (1995) and compute (4.1) if the conditional joint distribution of \(Z(t)\) and \(W(t - \Delta)\) given \([T^* \geq t]\) and \(x\) is assumed Gaussian and independent of \(x\) with \(Z(t)\) given by (3.7) and \(W(t)\) by (3.8). This may only be realistic in a low-risk situation, i.e. when the conditioning on \(T^* \geq t\) is negligible (Prentice, 1982), but it is useful for illustrating certain causes of attenuation; the effect of selective removal of high risk individuals is an additional potentially attenuating factor (see below). Letting

\[
a(t, \Delta) = \frac{\omega^2 + \sigma(t)\sigma(t - \Delta)\rho(t - \Delta, \Delta)}{\omega^2 + \sigma^2(t - \Delta) + \tau^2}
\]

the hazard derived from the additive model is then

\[
\alpha(t \mid W(t - \Delta), x) = \alpha_0(t) + \beta(t)\mu(t) + \beta(t)a(t, \Delta)(W(t - \Delta) - \mu(t - \Delta)) + \delta(t)^\prime x, 
\] (4.2)

while for the Cox model

\[
\alpha(t \mid W(t - \Delta), x) = \alpha_0^\ast(t, \Delta) \exp\left(\beta a(t, \Delta)(W(t - \Delta) - \mu(t - \Delta)) + \delta^\prime x\right). 
\] (4.3)

where

\[
\alpha_0^\ast(t, \Delta) = \alpha_0(t) \exp(\beta \mu(t)) \\
\times \exp\left(\frac{1}{2}\beta^2(\omega^2 + \sigma^2(t) + \frac{(\omega^2 + \sigma(t)\sigma(t - \Delta)\rho(t - \Delta, \Delta))^2}{\omega^2 + \sigma^2(t - \Delta) + \tau^2})\right). 
\] (4.4)

Several factors may contribute to attenuation. In the following we will discuss these, in turn, and illustrate them in Table 1 using simulations.
Table 1. Average estimated regression coefficients for prothrombin. BM: Brownian motion used for \( U_i(t) \) in Model (3.7), OU: Ornstein–Uhlenbeck process used for \( U_i(t) \). Updating is for each individual done at equally spaced time points. In the synchronous case, updating intervals are two years (before adjustment, see below), in the non-synchronous case intervals are uniformly distributed over one to three years. Small adjustments are carried out to obtain equal average length from measurements to events (about one year) in all situations. Trend implies a reduction in \( \mu(t) \) of five units per year. High risk is obtained with 500 individuals and the parameters of Section 3.3, while low risk is obtained with sample size 10000 and reduced basic hazard to keep the number of events constant. The true effect of prothrombin is \(-0.040\). Figures are based on 500 (low risk) or 2000 (high risk) simulations.

<table>
<thead>
<tr>
<th></th>
<th>BM</th>
<th>OU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Synchronous</td>
<td>Non-synchronous</td>
</tr>
<tr>
<td>No trend, low risk</td>
<td>-0.040</td>
<td>-0.039</td>
</tr>
<tr>
<td>No trend, high risk</td>
<td>-0.039</td>
<td>-0.039</td>
</tr>
<tr>
<td>Trend, low risk</td>
<td>-0.040</td>
<td>-0.038</td>
</tr>
<tr>
<td>Trend, high risk</td>
<td>-0.039</td>
<td>-0.037</td>
</tr>
</tbody>
</table>

Non-synchronous updating

For the Cox model, the ‘baseline hazard’ \( \alpha^*(t, \Delta) \) must not depend on any factor varying between individuals. This is achieved if \( \Delta \) is the same for all individuals, i.e. if updating is synchronous, otherwise \( \alpha^*(t, \Delta) \) will depend on \( \Delta \) as in the last factor in (4.4). For the additive model (4.2) this problem does not occur, but for both models the trend \( \mu(t - \Delta) \) subtracted from \( W(t - \Delta) \) in (4.2)–(4.3) will imply that the hazard function gets an extra variation between individuals under non-synchronous updating. The situation may also be seen as a missing covariate problem, suggesting that the likely effect of non-synchronous updating is an attenuation of regression coefficients. This was confirmed by simulations and Table 1 shows examples. Here, non-synchronicity is obtained by varying the updating intervals between individuals, but we also tested other non-synchronous updating patterns, such as equal updating intervals but shifted time 0, and updating with a certain probability. Substantial attenuation was only found in cases with a trend combined with marked variation in updating interval length.

Stochastic drift

Even with synchronous updating, for both of the hazard models considered the estimator obtained by replacing the true covariate \( Z(t) \) by the available value \( W(t - \Delta) \) will not estimate \( \beta \), but \( \beta \) times the factor

\[
a(t, \Delta) = \frac{\omega^2 + \sigma^2(t - \Delta)}{\omega^2 + \sigma^2(t - \Delta) + \tau^2} \frac{\omega^2 + \sigma(t)\sigma(t - \Delta)\rho(t, \Delta)}{\omega^2 + \sigma^2(t - \Delta)}.
\]

The first factor,

\[
v(t - \Delta) = \frac{\omega^2 + \sigma^2(t - \Delta)}{\omega^2 + \sigma^2(t - \Delta) + \tau^2},
\]

is a measurement error attenuation factor, similar to the factor seen in linear regression (Carroll et al., 1995). For the CSL1 parameter values estimated from the mixed linear model in Section 3.3 we obtain \( \hat{v} = \)
Attenuation caused by infrequently updated covariates in survival analysis

0.69 (note that for the stationary OU-process, \( \nu \) does not depend on time). The second factor represents the effect of ageing and is seen to depend on the correlation structure for \( Z(t) \). For a Brownian motion, the second factor is unity, that is, there will be no attenuation due to this factor alone. In this situation, \( E(U(t)|U(t-\Delta)) = U(t-\Delta) \) and the result may be seen as analogous to the result for the so-called Berkson measurement error model where for a true \( Z \) and an observed \( W \) we have \( E(Z|W) = W \) and no attenuation is predicted (Carroll et al., 1995, p. 150). However, for the OU-process the last factor of (4.5) decreases from 1 for \( \Delta = 0 \) towards \( \frac{\omega^2}{\omega^2+\sigma_u^2} \) according to the function \( \frac{\omega^2+\sigma_u^2 \exp(-\gamma \Delta)}{\omega^2+\sigma_u^2} \). For the CSL1 parameter values the attenuation will be pronounced if updating is infrequent, after two years the factor is 0.75. This effect is also seen in Table 1. Note especially the synchronous low-risk situation where, as expected, the Brownian motion case shows no attenuation while there is a substantial effect for the OU-case. Note also that Figure 1(a) corresponds to the synchronous OU-case without trend, and that this stochastic process with its stationary form is the dominating reason for the attenuation seen in the figure.

**Effects of selection**

Expressions (4.2)–(4.5) are all motivated by a low-risk assumption. To evaluate the effect of this assumption we first consider the situation where \( U(t) \) is a Brownian motion. Although \( E(U(t)|U(t-\Delta)) = U(t-\Delta) \) in this situation, the selective removal of high-risk individuals implies that the expected difference in covariate values for two survivors at some point \( t \) is less than the (observable) difference at time \( t-\Delta \), an effect that is quite marked when simulated with CSL1 parameters and mortality. The effect on parameter estimates is illustrated in Table 1. Attenuation is clearly present, although moderate. Only in situations more extreme than those quoted in Table 1, like the time-fixed BM case, was the effect substantial with an estimated coefficient of about –0.033 as opposed to the correct –0.040. One should note, however, that all covariate coefficients will be attenuated, thus in the latter simulations the estimated coefficient, \( \delta \), for age was about 0.041 as opposed to the correct 0.05. This may be seen as a consequence of the frailty or unobserved variable analogy mentioned in connection with (4.1), with \( U(t) - U(t-\Delta) \) as the unobserved variable.

In summary, stochastic development of the covariate may cause marked attenuation, but the size varies from no attenuation to marked attenuation depending on the form of the process. At least for situations approximating CSL1, this will be the main cause of attenuation due to ageing of covariates. However, combined with a trend, non-synchronous updating will add to the attenuation. Selective removal of high-risk individuals is a potential cause of attenuation, but in our simulations we found moderate differences between the low-risk and high-risk situations.

**5. Adjustment for attenuation**

In this section we will first review a number of existing methods for adjustment for attenuation due to ageing (and measurement errors) of covariates and illustrate some of these using the CSL1 data. The methods to be discussed are of two different types. The first one (Section 5.1) uses a two-step procedure where first some smoothing of the observed covariate values is carried out and afterwards the smoothed values are used in the estimation of the parameters in the hazard model. The second (Section 5.2) involves joint modelling of the covariate process and the failure times. In Section 5.3 we suggest simpler methods for adjustment inspired by the results of Section 4 and compare these with some of the two-step procedures using both the CSL1 trial and simulated data. All results regarding CSL1 are collected in Table 2 while the simulation study is reported in Table 3. An important conclusion is that while the existing methods appear to work well when many follow-up examinations are available they may be less appropriate with few updatings. In the final Section 5.4 we present a graphical method for assessing attenuation.
Table 2. CSL1 study: Estimated effect of prothrombin using a number of models for adjustment for ageing (and measurement errors). Adjustment for measurement errors is carried out using SIMEX, except for methods marked with (∗) which also take the effect of measurement errors into account

<table>
<thead>
<tr>
<th>Model</th>
<th>( \hat{\beta}_{\text{pro}} ) (SE)</th>
<th>Adjusted for measurement errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard time-dependent model</td>
<td>-0.025 (0.0025)</td>
<td>-0.036</td>
</tr>
<tr>
<td>Standard time-fixed model</td>
<td>-0.015 (0.0027)</td>
<td>-0.022</td>
</tr>
<tr>
<td>Two-step model: Stochastic coefficients</td>
<td>-0.034(∗)(0.0034)</td>
<td>-0.034(∗)</td>
</tr>
<tr>
<td>Conditional score</td>
<td>-0.038(∗)(0.0050)</td>
<td>-0.038(∗)</td>
</tr>
<tr>
<td>Time-dependent model (5.4):</td>
<td>-0.027 (0.0026)</td>
<td>-0.037</td>
</tr>
<tr>
<td>Time-fixed model (5.4):</td>
<td>-0.027 (0.0048)</td>
<td>-0.037</td>
</tr>
<tr>
<td>Deletion-extrapolation</td>
<td>-0.028 (0.0032)</td>
<td>-0.040</td>
</tr>
</tbody>
</table>

5.1 Two-step procedures: regression calibration

In regression calibration (see Carroll et al., 1995) the basic idea is to use estimates of expected values of the covariates given the observed variables in the estimation procedure. Smoothing of the observed covariate values may be seen as one way of obtaining such estimates, and thus a way of avoiding or reducing attenuation due to both measurement errors and ageing. Tsiatis et al. (1995) used growth curve models for this purpose while Raboud et al. (1993) used several parametric and non-parametric smoothing techniques. Dafni and Tsiatis (1998) assumed a linear random coefficients regression model for \( Z_i(t) \) and an additive measurement error model like (3.8). Boscardin et al. (1998) and Bycott and Taylor (1998) studied stochastic models for CD4 counts and their effect on AIDS incidence. The reported results indicate that models adapted to specific problems may function well, although the results of Tsiatis and Davidian (2001) indicate that this is not always the case.

When selecting a model for a specific situation, one important factor affecting the choice is the number of updatings. For data with few updatings for a large fraction of the individuals, as is the case for CSL1, one has to select parsimonious smoothing models. We thus tested a two-step procedure with a linear random coefficients model \( Z_i(t) = A_{0i} + A_{1i}t \) for prothrombin. An estimate of -0.034 (0.0034) was obtained, representing a marked de-attenuation compared to the `naive' estimate of -0.025 obtained using the last recorded value prior to \( t \) (cf. Table 2).

5.2 Joint modelling of covariates and failure time

Probably the best method to adjust for both measurement errors and ageing is jointly modelling the development over time of the covariate process and the hazard function given the covariate history. For the additive model with a quadratic effect of \( Z_i(t) \), the Manton–Woodbury model (see Woodbury and Manton, 1977; Martinussen and Keiding, 1997), allows explicit computations. DeGruttola and Tu (1994) studied a model where the covariate process and the survival time (the latter via an accelerated failure time type model) are driven by the same random factors. For the Cox model, Wolfssohn and Tsiatis (1997) used Gaussian growth models and the EM algorithm to jointly estimate all model parameters. Henderson et al. (2000) and Xu and Zeger (2001) generalized this approach. Faucett and Thomas (1996) used Markov chain Monte Carlo in a fully parametric model with a piecewise constant baseline hazard.

Finally, Tsiatis and Davidian (2001) studied the model (3.6) with \( Z_i(t) \) given by a linear random coefficients regression model, with \( Z_i(t) = A_{0i} + A_{1i}t \) and \( W_i(t) \) given by (3.8). For individuals with at least two measurements for \( W_i(\cdot) \) they derived a conditional score given a series of sufficient statistics.
Attenuation caused by infrequently updated covariates in survival analysis

\(S_i(t)\) for \((A_{0i}, A_{1i})\) depending on data for individual \(i\) from \((0, t)\). This conditional score then only depends on the parameters in (3.6) and on the measurement error variance \(\tau^2\). Large sample properties for the resulting estimators \(\hat{\beta}\) and \(\hat{\delta}\) were assessed for known \(\tau^2\).

When this latter method was tested on CSL1, we obtained a coefficient of \(-0.038 (0.0050)\) for prothrombin, a value clearly further away from the naive estimate than the one obtained with the two-step procedures. This is consistent with the results of Tsiatis and Davidian (2001), where this method compensated for attenuation efficiently and more effectively than the two-step procedures. The method thus appears as a good alternative, although the estimates have a somewhat high standard error and at least two measurements are needed on each individual. For CSL1 this was fulfilled for only 417 of the 488 patients.

5.3 A simple approach: adjusted covariates

If several measurements are available for each individual the methods above, and in particular the one of Tsiatis and Davidian (2001), provide good alternatives. Many of the methods are, however, fairly complicated and also seem less appropriate if few updatings are available on each individual. We have therefore evaluated a simple alternative that requires few updatings (and may even be used without any updating).

The regression calibration idea may be used to motivate the approach. We then use results like those of Section 4 to obtain approximate expected values at some time \(t\) based on the last measurement before \(t\). To illustrate, first assume that updating is synchronous or there is no marked trend. Equation (4.3) then suggests using an adjusted covariate of the form

\[a(t, \Delta)(W(t - \Delta) - \mu).\]  

(5.1)

If we neglect measurement errors and assume an OU-process then the first factor of \(a(t, \Delta)\) in (4.5) is \(v(t - \Delta) = 1\) and the second factor has the form \(a + be^{-c\Delta}\). In this case, (5.1) suggests entering terms of the form

\[\beta_1 W(t - \Delta) + \beta_2 W(t - \Delta)e^{-c\Delta} + \beta_3 e^{-c\Delta}\]  

(5.2)

in the regression model where the true effect, \(\beta\), corresponds to \(\Delta = 0\), i.e. \(\hat{\beta} = \hat{\beta}_1 + \hat{\beta}_2\).

This form was derived to compensate for the effect of the stochastic development which in our studies appeared as the main cause of attenuation. However, since (5.2) already includes several parameters and since additional terms included to adjust for factors like selective survival or trends will be correlated to those in (5.2), we tested how well (5.2) worked. The parameter \(c\) cannot be estimated by maximizing the standard partial likelihood (3.2), but we may use a profile likelihood where we maximize for a suitable set of fixed values of \(c\) and use the results corresponding to the maximal likelihood value.

Since even (5.2) includes many parameters and requires a profile likelihood, we have evaluated simpler functions. One possibility is to select a fixed value for \(c\) in (5.2). Another alternative is to use a linear, a piecewise linear or a piecewise constant function of \(\Delta\) for \(a(t, \Delta)\) in (5.1). For the simple linear case we get terms of the form

\[\beta_1 W(t - \Delta) + \beta_2 W(t - \Delta)\Delta + \beta_3 \Delta,\]  

(5.3)

and for the piecewise constant function with two intervals we have

\[\beta_1 W(t - \Delta) + \beta_2 W(t - \Delta)I(\Delta > L) + \beta_3 I(\Delta > L),\]  

(5.4)

where \(L\) is a chosen limit and \(I(\cdot)\) is an indicator function. In both (5.3) and (5.4), \(\hat{\beta} = \hat{\beta}_1\).
Also when we only have a single initial measurement, but a covariate that evolves with time, similar terms may be introduced in the Cox model. The situation is technically simpler since measurements are trivially synchronous. Common trends are thus uninteresting, and for instance (5.2) reduces to \( \beta_1 W(0) + \beta_2 W(0)e^{-ct} \) and (5.4) to \( \beta_1 W(0) + \beta_2 W(0)I(t > L) \).

When applying the methods on CSL1, we had to take into account that several measurements were unscheduled and often taken from patients in ‘crisis’, typically close to death (see Liestøl and Andersen, 2002). Such observations will invalidate models like (5.2). Model (5.4) is, however, robust in this respect. When used on CSL1 with \( L = 1 \) year, the coefficient for prothrombin was \(-0.027 (0.0026)\) when all updatings were used and \(-0.027 (0.0048)\) when only the initial measurement was used. The modest degree of de-attenuation is a consequence of the fairly frequent updating of CSL1. Since this method only compensates for ageing, measurement errors must be accounted for separately. The SIMEX method (see Carroll et al, 1995) using the value \( \hat{\tau}_2 = 213 \) (Section 3.3) results in a coefficient of \(-0.037\) for prothrombin.

To evaluate these methods more systematically in situations with few updatings, we used simulations mimicking CSL1 but with infrequent and non-synchronous updating. Table 3 shows simulation results with updating every sixth month with probability 0.25, resulting in an average of slightly below three measurements for each individual. The table displays results for Models (5.2) and (5.4) both when all measurements are used and when only the initial measurements are used. For comparison we included in the study the two-step procedure using a random coefficient model for the covariate. We also tested the simple interpolation technique described by, e.g. Altman and Stavola (1994) and Collett (1994), where \( Z_i(t) \) at some point \( t \) is obtained by interpolation from observations before and after \( t \) (if available). Finally, we included an alternative where one filters out individuals temporarily if they have a value of \( \Delta \) larger than some chosen cut-off value (e.g. Gail, 1981; Raboud et al., 1993; Andersen et al., 1993, Section III.4). We have simulated both with and without measurement errors and both with an OU-process and a Brownian motion; the latter is hardly a realistic model in typical biomedical studies and should be regarded as a robustness check.

Based on the simulation results in Table 3, we note

- The exponentially decreasing model (5.2) compensates well on average for attenuation. However, the variance is high and the maximum for \( c \) in the profile likelihood was often at the boundary of the allowed interval. Thus, estimation with a model of this complexity requires large sample sizes.
- As expected, the coefficient obtained for the first interval using the piecewise constant model (5.4) is attenuated. The attenuation is, however, moderate and sample size requirements are clearly lower than for model (5.2).
- As expected, artificial censoring when time since measurements exceeds some limit gives estimates close to those obtained for the first interval with piecewise constant functions. Estimating with a piecewise constant function (Model (5.4)) utilizes more of the data and provides extra information on the attenuation at a longer distance from updating, and will probably in most cases be a better approach.
- The two-step procedure using a linear random coefficient model aims at compensating both for ageing and measurement errors, and in the simulated situations achieves this with only moderate attenuation. When a trend is present, the result depends on how one handles observations with single measurements. In Table 3, we have extrapolated from the estimate at time 0 using the estimated common trend.
- The simulations show that the two-step procedure with linear interpolation should not be recommended, particularly when a trend is present.
Table 3. Simulation study: Estimated effect of prothrombin for the models of Section 3.3 when updating frequency is low. The entries correspond to the average estimated regression coefficient for prothrombin, the true effect being $-0.040$. Entries in parentheses give the standard deviation for the coefficients in the 1000 simulations used as basis for the table. Simulations are carried out with an Ornstein–Uhlenbeck (OU) process or a Brownian motion (BM) for the stochastic drift $U_i(t)$. Updatings for the time-dependent models are carried out every half-year with probability $0.25$. The trend implies a reduction in prothrombin values of five units per year.

<table>
<thead>
<tr>
<th></th>
<th>No measurement errors</th>
<th>With measurement errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OU</td>
<td>OU with trend</td>
</tr>
<tr>
<td>Continuous updating</td>
<td>$-0.040$</td>
<td>$-0.040$</td>
</tr>
<tr>
<td>Standard time-</td>
<td>$-0.033$</td>
<td>$-0.030$</td>
</tr>
<tr>
<td>dependent model</td>
<td>(0.0030)</td>
<td>(0.0027)</td>
</tr>
<tr>
<td>Exponential model (5.2)</td>
<td>$-0.041$</td>
<td>$-0.040$</td>
</tr>
<tr>
<td></td>
<td>(0.0049)</td>
<td>(0.0050)</td>
</tr>
<tr>
<td>Piecewise</td>
<td>$-0.037$</td>
<td>$-0.036$</td>
</tr>
<tr>
<td>constant model (5.4)</td>
<td>(0.0039)</td>
<td>(0.0037)</td>
</tr>
<tr>
<td>Artificially</td>
<td>$-0.037$</td>
<td>$-0.037$</td>
</tr>
<tr>
<td>censored model</td>
<td>(0.0040)</td>
<td>(0.0038)</td>
</tr>
<tr>
<td>Two-step model:</td>
<td>$-0.035$</td>
<td>$-0.024$</td>
</tr>
<tr>
<td>Linear interpolation</td>
<td>(0.0030)</td>
<td>(0.0025)</td>
</tr>
<tr>
<td>Two-step model:</td>
<td>$-0.039$</td>
<td>$-0.038$</td>
</tr>
<tr>
<td>Stochastic coefficients</td>
<td>(0.0039)</td>
<td>(0.0038)</td>
</tr>
<tr>
<td>Standard time-</td>
<td>$-0.026$</td>
<td>$-0.024$</td>
</tr>
<tr>
<td>fixed model</td>
<td>(0.0032)</td>
<td>(0.0029)</td>
</tr>
<tr>
<td>Exponential model (5.2)</td>
<td>$-0.041$</td>
<td>$-0.042$</td>
</tr>
<tr>
<td></td>
<td>(0.0067)</td>
<td>(0.0066)</td>
</tr>
<tr>
<td>Piecewise</td>
<td>$-0.037$</td>
<td>$-0.037$</td>
</tr>
<tr>
<td>constant model</td>
<td>(0.0050)</td>
<td>(0.0043)</td>
</tr>
</tbody>
</table>

In conclusion, methods based on covariate adjustment seem useful when few updatings are available, but except in cases with large sample sizes one will have to use simple models like (5.3) or (5.4). With increasing numbers of updatings for each individual, methods like the two-step procedure using a linear random coefficients model becomes more attractive, and with a sufficient number of updatings the method of Tsiatis and Davidian (2001) seems preferable. When choosing a method, we also have to take into account whether adjustment for measurement errors is warranted, i.e. whether we aim at the ‘real’ effect of the covariate or at the predictive value, cf. Section 1. However, if an estimate of the measurement error variance is available, either from repeated measurements or from validation data, we may switch between the two kinds of estimates; a rough way of doing this is to use the attenuation factor $\nu(\cdot)$, see (4.6).

We note that models similar to (5.2) have been used in other situations. Thus, when assessing the effect of transplantation on survival, Cox and Oakes (1984) used profile likelihoods to estimate with covariates of the form $\beta_1 I_i(t) + \beta_2 T_i(t)e^{-\Delta_i}$, where $\Delta_i$ is the time since transplantation and $I_i(\cdot)$ is an indicator.
that changes from 0 to 1 when transplantation takes place for individual \(i\). Jagger and Sutton (1991) used a similar model when assessing mortality following bereavement DeBruijne et al. (2001) adjusted for time elapsed since the last measurement in a study of survival prediction. In the time-fixed situation, the covariates corresponding to the linear model (5.3) will be \(\beta_1 W(0) + \beta_2 W(0) \cdot t\), a form often used in tests of the proportional hazards assumption.

5.4 Graphical assessment of attenuation: a 'DELEX' algorithm

To visualize the attenuation caused by the ageing of measurements for data sets like CSL1, we may use an explorative method that in some respects resembles the SIMEX procedure (see Carroll et al., 1995) for correcting for measurement errors. The basic idea is to estimate the regression coefficient using different subsets of the follow-up measurements, and plot the coefficients versus some measure of how old the measurements in average are when used in the estimation. Such plots will clearly be exploratory tools: it is not obvious how best to select the sample of measurement subsets or to quantify the collective ‘age’ of measurements. However, based on simulations we consider the plots informative despite these ambiguities.

Two points on such plots are naturally obtained by using, respectively, only the initial measurement in a time-fixed model and by using all measurements in a time-dependent model. To get additional points on the plot, we may use systematic deletion of measurements as in Figure 1(a). However, especially with small datasets it appears more appropriate to use repeated random deletions so that all measurements are used in different combinations. The random deletions imply that we get a mix of short and long intervals between measurements, but the effect of this appears to be moderate: when we remade the curve for the large dataset in Figure 1(a) by completely random deletions, the obtained curve was close to the one obtained with systematic deletion. As a measure of the average age of measurements, we computed for each event the time distance back to the preceding measurement for the patient experiencing the event, and then used the mean of these distances.

In a Cox model for CSL1 with prothrombin and age as covariates, the coefficient for prothrombin is \(-0.015\) for the time-fixed model and \(-0.025\) for the time-dependent model using all measurements (Section 2 and Table 2). The corresponding mean ages of measurements are 2.9 years in the time-fixed case and 0.4 years in the time-dependent case. The remaining crosses in Figure 1(b) are obtained as averages over 100 datasets with random deletions of 20%, 40%, 60%, 80% and 90% of the follow-up measurements. The results of Section 4 suggest that a curve of the form \(a + be^{-ct}\) should be fitted to these points. However, simulations (see below) indicated that a second-order polynomial provides more stable results and the curve in Figure 1(b) is of this form.

Fitting a curve as in Figure 1(b) provides predictions of what would happen with other updating frequencies, including continuous updating where the procedure leads to an estimate of \(-0.028\) (Table 2). Simulations indicate that this ‘deletion-extrapolation’ procedure will normally provide a fair picture of what happens if extrapolation distance is moderate, at least in situations similar to our CSL1 model. For instance, in 100 simulations using the standard OU-model and follow-up measurements every half-year with probability 0.5, the estimated coefficient always using the last measurement was \(-0.036\) (standard deviation 0.0026) while graphical extrapolation to continuous updating using a second-order polynomial gave a value close to the correct \(-0.04\) (0.0031). On average, extrapolation using \(a + be^{-ct}\) gave similar results but in a few cases extreme and unreliable results were obtained. We therefore recommend using a second order polynomial for extrapolation.

Artificial temporary censoring (or ‘filtering’: see Andersen et al., 1993, Section III.4) of observations is an alternative way of using subsets of the data. By reducing the maximal distance from measurements to events, this creates datasets with shorter average age of measurements. Such censoring may either be used alone as a method to obtain estimates closer to continuous updating, or as a check of the method above.
The ‘C’ in Figure 1(b) has been obtained by censoring observations one year after each measurement and re-entering the patient after the next measurement using delayed entry (see Section 3.1). Some events will be eliminated by this procedure so that estimation uncertainty increases.

When estimates of measurement errors exist, we may assess their effect and illustrate this on the graph. In Figure 1(b) we have done this using the estimate (\( \hat{\nu} = 0.69 \)) of the attenuation factor from Section 4 and checking the result by SIMEX for the full time-dependent model and the time-fixed model. The adjusted estimate for continuous updating is \(-0.040\) (Table 2).

Although we advocate DELEX as an exploratory rather than an inferential tool it may be of interest to obtain standard errors of estimates derived using this procedure. A natural route to follow would be to use the bootstrap (Carroll et al., 1995) though the actual computations may become somewhat involved. A bootstrap (using only 250 replications) gave a standard error of 0.0032 (Table 2). The same value was obtained as standard error for the intercept when fitting the second order polynomial in Figure 1(b) using the inverse of the estimated variances as weights.

6. FINAL COMMENTS

Our discussion has focused on the attenuation of regression coefficients caused by ageing of covariates. As revealed by the discussion in Section 4, the cause of the attenuation may be a combination of stochastic drift, trends and non-synchronous updating. With small data sets one will often not be able to estimate the size of this attenuation. However, it may still be useful to assess whether non-negligible effects are likely, for instance by using simple simulation models and examining the attenuation seen under different assumptions. With larger data sets one may actually try to compensate for the attenuation or predict what would happen under other sampling schemes. One set of methods uses two steps: first the updated covariates are smoothed and then the smoothed values are used in the regression. Another possibility is simultaneous modelling of covariate development and survival, and in this context the method of Tsiatis and Davidian (2001) seems especially promising, provided a sufficient number of updatings is available. Both of these sets of methods take possible measurement errors into account. In Section 5 we have studied the simpler possibility of using adjusted covariates, which has the advantage that the method may be used with few (or no) updatings. We also considered a graphical (‘DELEX’) procedure. Neither of these techniques requires modelling of the time-dependent covariate. The evaluated methods worked satisfactorily in the situations considered, but we strongly advocate the use of various forms of sensitivity analysis by manipulation of the data (e.g. by artificial censoring) and testing of alternative models. It should be emphasized that for the simpler methods suggested, correction for measurement errors must be carried out separately and that evaluation of the uncertainty of the corrected estimates may not be straightforward.

In conclusion, regression models for survival data with time-dependent covariates are in many respects considerably more involved than models which only include time-fixed covariates. Updated covariates will typically be rich in information and allow assessment of various effects. The associated cost is that the analyses will often not end with a single coefficient describing the effect of the covariate, but with a discussion supported by a set of estimates.

ACKNOWLEDGEMENTS

This research was partially supported by a grant (R01 CA54706-07) from the National Cancer Institute. The activities of the Danish Epidemiology Science Centre are supported by a grant from the Danish National Research Foundation. Comments on an earlier version of the manuscript from Ørnulf Borgan and Niels Keiding are greatly appreciated. We are grateful to A. A. Tsiatis and Xiao Song for making a computer program for estimation using the conditional score approach available to us.
REFERENCES


Attenuation caused by infrequently updated covariates in survival analysis


[Received November 2, 2001; first revision October 14, 2002; second revision April 3, 2003; accepted for publication May 7, 2003]