Let $A.(t) = \sum_{j=1}^{J} A_j(t)$. Since the outcome of $J$ does not affect the censoring of $X$ and only affects the $G$ part of the likelihood, it is not hard to show that the orthogonal complement of the nuisance tangent space for $\mu$ is identical to the orthogonal complement of the nuisance tangent space class of all estimating functions for $\mu$ for the reduced data structure $(C, \bar{X}(C))$, as presented in this chapter. The corresponding estimating functions only require an estimate of the intensity $\alpha(t) = E(dA.(t) \mid \mathcal{F}(t))$ (i.e., of the conditional distribution of $C$, given $X$) and, in particular, the survivor function

$$\bar{G}(t \mid X) \equiv P(C \geq t \mid X) = \exp\left(-\int_{[0,t)} \alpha(s)ds\right) \text{ of } C, \text{ given } X.$$

This censoring mechanism could be fit with the reduced data structure $(C, \bar{X}(C))$. However, if knowledge is available on the cause specific intensities of $A_j$, then a natural framework for fitting a continuous intensity $\alpha$ of $A$ w.r.t. $\mathcal{F}(t)$ is to assume multiplicative intensity models

$$\alpha_j(t) \equiv E(dA_j(t) \mid \mathcal{F}(t)) = I(\tilde{T} \geq t)\lambda_{0j}(t)\exp(\alpha_j W_j(t))$$

and use $\alpha(t) = \sum_{j=1}^{J} \alpha_j(t)$. Again, one can use the S-plus function Coxph() to fit the intensities $\alpha_j$, $j = 1, \ldots, J$ (Andersen, Borgan, Gill, and Keiding, 1993). Similarly, such a strategy could be applied if the censoring mechanism is discrete. In this case we note that $\bar{G}(t \mid X) = \prod_{[0,t)} \prod_{j=1}^{J}(1-\lambda_j(t))$, where $\lambda_j(t) \equiv E(dA_j(t) \mid A_1(t) = \ldots = A_{j-1}(t) = 0, C \geq t, \mathcal{F}(t))$. Thus, following Robins (1993a), we conclude that the methods presented in this chapter for single cause censoring can be applied by only using the additional data $J$ to obtain an estimate $G_n$ of $G(t \mid X)$ but further ignoring $J$.

## 3.2 Examples

### *3.2.1 Right-censored data on a survival time*

Consider a longitudinal study in which the outcome of interest is the survival time $T$. We assume that each subject is regularly followed up and relevant time-dependent measurements are taken at each visit. Let $R(t) = I(T > t)$ be the survival status of the subject at time $t$. Let $L(t)$ be the time-dependent process representing these measurements at time $t$. We denote the history of $L(t)$ with $\bar{L}(t) = \{L(s) : s \leq t\}$. The time-independent baseline covariates are included in the vector $L(0)$. For full data on a subject we have $X = (T, \bar{L}(T)) = (\bar{R}(T), \bar{L}(T))$. Let $C$ be the dropout time of the subject. We observe each subject up to the minimum of $T$ and $C$. Thus we observe $n$ i.i.d. observations $Y_1, \ldots, Y_n$ of

$$Y = \Phi(T, \bar{L}(T), C) \equiv (\tilde{T} = T \wedge C, \Delta = I(T \leq C), \bar{L}(\tilde{T})).$$



One can also represent this data structure as

$$Y = (\tilde{T}, \Delta, \bar{X}(\tilde{T}))$$

since $X(t) = (R(t), L(t))$. Thus, this data structure is of the type considered in this chapter. Important parameters might be 1) the marginal survival function of $T$, 2) regression parameters in a generalized linear regression model of $\log(T)$, onto some baseline covariates and/or treatment or regression parameters and 3) regression parameters of a Cox proportional hazards model for the hazard of $T$, adjusting for some baseline covariates, treatment, and possibly some of the time-dependent covariates.

Estimation with this data structure is a complicated and practically important problem. This data structure is an important extension of the marginal right censored data structure on $T$. For an extensive description of the literature on the marginal univariate right-censored data structure, we refer the reader to Andersen, Borgan, Gill and Keiding (1993). For maximum likelihood (or more general partial likelihood) estimation and inference with multiplicative intensity models such as the Cox proportional hazards model for this data structure, see Andersen, Borgan, Gill and Keiding (1993). Such models are of low enough dimension that the maximum likelihood estimator can be used. However, if one is interested in more marginal parameters, such as the marginal distribution of $T$ or a regression model of $T$ on a subset of the baseline and/or time-dependent covariates, then this literature does not provide an appropriate methodology.

Locally efficient estimation of regression parameters in Cox proportional hazards and accelerated failure time models for this data structure, allowing covariates outside the model, has been studied in Robins (1993a) and Robins and Rotnitzky (1992). The importance of assuming a CAR model in these real-life applications has been argued in detail in these papers. Locally efficient estimation of the marginal distribution of $T$ has been studied and implemented with simulations in Hubbard, van der Laan and Robins (1999). Robins (1996) studies locally efficient estimation in the median regression model. Robins and Finkelstein (2000) apply the inverse probability of censoring weighted estimator of survival to analyze an AIDS clinical trial.

### *3.2.2 Right-censored data on quality-adjusted survival time*

Consider a longitudinal study in which a quality adjusted survival time of a subject is the time variable of interest. Let $T$ be the chronological survival time of the subject. Let $V(t)$ be the state of health of the subject at time $t$. Typically, one assumes that the state space is finite so that $V(t) \in \{1, \ldots, k\}$ for some $k$, but this is not necessary. We define the quality adjusted lifetime as $U \equiv \int_0^T Q(V(t))dt$, where $Q$ is some given function. One sensible definition of quality-adjusted lifetime can be obtained by defining $V(t)$ as the quality of life at time $t$ on a scale between 0 and 1 and $Q(t) = t$. In this case, one might define $V(t) = 1$ if the subject