HHS Public Access

Regression Models and Multivariate Life Tables

Ross L. Prentice,
Public Health Sciences Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle, Washington, USA 98109

Shanshan Zhao
National Institute of Environmental Health Sciences, 111 TW Alexander Dr, Rall Building, Research Triangle Park, North Carolina, USA 27709

Abstract

Semiparametric, multiplicative-form regression models are specified for marginal single and double failure hazard rates for the regression analysis of multivariate failure time data. Cox-type estimating functions are specified for single and double failure hazard ratio parameter estimation, and corresponding Aalen–Breslow estimators are specified for baseline hazard rates. Generalization to allow classification of failure times into a smaller set of failure types, with failures of the same type having common baseline hazard functions, is also included. Asymptotic distribution theory arises by generalization of the marginal single failure hazard rate estimation results of Danyu Lin, L.J. Wei and colleagues. The Péano series representation for the bivariate survival function in terms of corresponding marginal single and double failure hazard rates leads to novel estimators for pairwise bivariate survival functions and pairwise dependency functions, at specified covariate history. Related asymptotic distribution theory follows from that for the marginal single and double failure hazard rates and the continuity, compact differentiability of the Péano series transformation and bootstrap applicability. Simulation evaluation of the proposed estimation procedures are presented, and an application to multiple clinical outcomes in the Women's Health Initiative Dietary Modification Trial is provided. Higher dimensional marginal hazard rate regression modeling is briefly mentioned.

Keywords

Bivariate survival function; Composite outcome; Cross ratio; Empirical process; Hazard rates; Marginal modeling; Multivariate failure times

1 Introduction

Sir David Cox's landmark paper (Cox, 1972) revolutionized the methods for analyzing censored failure time regression data. Cox regression, along with Kaplan–Meier (KM) survival function estimators, quickly became core methods for the analysis of univariate failure time data.

In the subsequent 45 years a considerable statistical literature has arisen proposing methods that are built upon univariate Cox regression for the analysis of multivariate failure time regression data. An important contribution was provided by Andersen and Gill (1982), who

used the same semiparametric exponential model form for the counting process intensity, which models failure rates for multivariate events on a single failure time axis conditional on all preceding failure, censoring and covariate information. As such these methods are suited to examining hazard ratio dependencies on covariates after allowing for the preceding failure history for the correlated set of outcomes, but cannot be used to examine covariate effects on hazard rates without the counting process conditioning. Frailty models (e.g., Andersen et al., 1993; Hougaard, 2000; Aalen et al., 2010; Duchateau and Janssen, 2010; Wienke, 2011) have also played a major role in multivariate failure time data analysis methods. These models avoid intensity models that depend explicitly on the individual's preceding failure time counting process by assuming independence between the failure times given a random effect, or frailty variable, which is typically assumed to act multiplicatively on the hazard rates for the correlated failure times given preceding covariate histories. These modeling methods are suited to the study of dependencies or clustering among correlated failure times given preceding covariate histories, but are less suited to the study of hazard rate associations with covariate histories themselves. In particular, marginal hazard rates given covariates that are induced by frailty models typically do not reduce to the standard form of the Cox models for marginal single failure hazard rates.

The copula model approach (e.g. Clayton, 1978; Oakes, 1986, 1989; Nan et al., 2006; Nelsen, 2007; Bandeen-Roche and Ning, 2008; Hu et al., 2011) to multivariate failure time modeling avoids these issues by assuming the standard Cox models for marginal single failures hazard rates, and by bringing together corresponding marginal survival functions (given covariates) through a copula function having a low dimensional parameter that controls dependency. A two-stage data analysis (e.g. Shih and Louis, 1995) then retains the usual marginal single failure hazard parameter rate estimators, while also providing estimators of copula distribution parameters. If the assumed copula model is a good fit to the data, this approach can provide a simple parametric description of dependencies among failure times, given covariates. This description can allow such dependencies to depend on baseline covariates, but the copula approach is not suited to estimating dependencies that are functions of covariates that evolve over the study follow-up period(s).

Frailty and copula models typically embrace a limited class of dependencies among the multivariate failure times. A semiparametric marginal regression modeling approach can add valuable analytic flexibility. Importantly the modeling approach that is considered here includes semiparametric regression models for both marginal single and marginal double failure hazard rates, and has the potential to add readily interpretable information on regression influences on failure time outcomes jointly, beyond that from Cox model analyses of the univariate outcomes. Asymptotic distribution theory for estimators of the regression parameters and corresponding baseline rates in the marginal single and double failure rate models will be developed by generalizing the empirical process results of Spiekerman and Lin (1998) for marginal single failure hazard rates and Lin et al. (2000) for recurrent event data, to embrace both marginal single and double failure rate models. These methods are largely complementary to well-developed semiparametric regression methods for the failure counting process intensity that considers the regression associations with the rates that condition on the entire preceding failure time counting process history for the correlated set of outcomes. The methods that are presented here aim to elucidate population-averaged

regression effects, in contrast to the subject-specific associations targeted by intensity models.

2 Bivariate Failure Time Regression Modeling and Estimation

2.1 Marginal single and double failure rate regression

Consider bivariate failure times $T_1 > 0$ and $T_2 > 0$ that are subject to right censoring by C_1 0 and C_2 0, respectively, with the usual convention that failures precede censorings in the event of tied times. Also suppose that the pair (T_1, T_2) is accompanied by a bivariate covariate that may stochastically evolve over the study follow-up period. Let $z(t_1, t_2)$ denote a vector of measured covariates at follow-up times (t_1, t_2) and let $Z(t_1, t_2) = z(0, 0) \vee \{z(s_1, s_2); s_1 < t_1, s_2 < t_2\}$ the covariate history prior to (t_1, t_2), where '$\vee$' denotes composition. One can define the marginal single failure hazard rate processes, $\Lambda_{10}\{\,\cdot\,, 0; Z(\,\cdot\,, 0)\}$ and $\Lambda_{01}\{0, \,\cdot\,; Z(0, \,\cdot\,)\}$ by

$$\Lambda_{10}\{dt_1, 0; Z(t_1, 0)\} = P\{T_1 \in [t_1, t_1 + dt_1); T_1 \geq t_1, Z(t_1, 0)\} \text{ for } t_1 \geq 0, \text{ and}$$
$$\Lambda_{01}\{0, dt_2; Z(0, t_2)\} = P\{T_2 \in [t_2, t_2 + dt_2); T_2 \geq t_2, Z(0, t_2)\}, \text{ for } t_2 \geq 0,$$

and marginal double failure hazard rate process $\Lambda_{11}\{\,\cdot\,, \,\cdot\,; Z(\,\cdot\,, \,\cdot\,)\}$ by

$$\Lambda_{11}\{dt_1, dt_2; Z(t_1, t_2)\} = P\{T_1 \in [t_1, t_1 + dt_1), T_2 \in [t_2, t_2 + dt_2); T_1 \geq t_1, T_2 \geq t_2, Z(t_1, t_2)\},$$

for all t_1 0 and t_2 0. An independent censoring assumption (given Z) for estimation of parameters in these hazard rate processes requires that lack of censoring in $[0, t_1)$, in $[0, t_2)$, and in $[0, t_1) \times [0, t_2)$ can be added, respectively, to the conditioning events in these expressions without altering the failure rates, for any (t_1, t_2).

Though a variety of regression models could be entertained for these marginal single and double failure hazard rates, we will focus on Cox-type semiparametric models, and write

$$\Lambda_{10}\{dt_1, 0; Z(t_1, 0)\} = \Lambda_{10}(dt_1, 0)\exp\{X(t_1, 0)\beta_{10}\}, \tag{1}$$

$$\Lambda_{01}\{0, dt_2; Z(0, t_2)\} = \Lambda_{01}(0, dt_2)\exp\{X(0, t_2)\beta_{01}\}, \text{ and} \tag{2}$$

$$\Lambda_{11}\{dt_1, dt_2; Z(t_1, t_2)\} = \Lambda_{11}(dt_1, dt_2)\exp\{X(t_1, t_2)\beta_{11}\}. \tag{3}$$

Here $\Lambda_{10}(\,\cdot\,, 0), \Lambda_{01}(0, \,\cdot\,)$, and $\Lambda_{11}(\,\cdot\,, \,\cdot\,)$ are unspecified 'baseline' hazard functions at zero values for the corresponding regression variables $X(t_1, 0) = \{X_1(t_1, 0), X_2(t_1, 0), \dots\}, X(0, t_2) = \{X_1(0, t_2), X_2(0, t_2), \dots\},$ and $X(t_1, t_2) = \{X_1(t_1, t_2), X_2(t_1, t_2), \dots\}$. These regression variables, with sample paths that are continuous from the left with limits from the right, are each fixed length vectors (i.e., same length vectors for all

study subjects at all times) formed from $\{t_1, Z(t_1, 0)\}$, $\{t_2, Z(0, t_2)\}$ and $\{t_1, t_2; Z(t_1, t_2)\}$ respectively. Also, β_{10}, β_{01} and β_{11} are corresponding (column vector) regression parameters to be estimated. The hazard ratio factors on the right side of (1)–(3) aim to quantify the dependence of these failure rates on the pertinent preceding covariate history.

Consider a random sample
$S_{1i} = T_{1i} \wedge C_{1i}, \delta_{1i} = I[T_{1i} = S_{1i}], S_{2i} = T_{2i} \wedge C_{2i}, \delta_{2i} = I[T_{2i} = S_{2i}], Z(S_{1i}, S_{2i}), \text{ for } i = 1, \dots, n$
from a study cohort, where $\wedge$ denotes minimum and $I[\cdot]$ is an indicator function. These observations define corresponding counting processes N_{1i}, N_{2i} and 'at risk' processes Y_{1i}, Y_{2i} via

$$N_{1i}(dt_1) = \begin{cases} 1 \text{ if } S_{1i} = t_1 \text{ and } \delta_{1i} = 1 \\ 0 \text{ otherwise} \end{cases} ; N_{2i}(dt_2) = \begin{cases} 1 \text{ if } S_{2i} = t_2 \text{ and } \delta_{2i} = 1, \\ 0 \text{ otherwise} \end{cases}$$

$$Y_{1i}(t_1) = \begin{cases} 1 \text{ if } S_{1i} \geq t_1 \\ 0 \text{ otherwise} \end{cases} ; \text{ and } Y_{2i}(t_2) = \begin{cases} 1 \text{ if } S_{2i} \geq t_2, \\ 0 \text{ otherwise} \end{cases}$$

for $i = 1, \dots, n$. From the above expressions one can define processes $L_{10i}(\cdot, 0; \beta_{10}), L_{01i}(0, \cdot; \beta_{01})$ and $L_{11i}(\cdot, \cdot; \beta_{11})$ that have zero means under (1)–(3) respectively, by

$$L_{10i}(dt_1, 0; \beta_{10}) = N_{1i}(dt_1) - Y_{1i}(t_1)\Lambda_{10}(dt_1, 0)\exp\{X_i(t_1, 0)\beta_{10}\},$$
$$L_{01i}(0, dt_2; \beta_{01}) = N_{2i}(dt_2) - Y_{2i}(t_2)\Lambda_{01}(0, dt_2)\exp\{X_i(0, t_2)\beta_{01}\}, \text{ and}$$
$$L_{11i}(dt_1, dt_2; \beta_{11}) = N_{1i}(dt_1)N_{2i}(dt_2) - Y_{1i}(t_1)Y_{2i}(t_2)\Lambda_{11}(dt_1, dt_2)\exp\{X_i(t_1, t_2)\beta_{11}\}.$$

An estimating equation for the hazard ratio parameter $\beta = (\beta'_{10}, \beta'_{01}, \beta'_{11})'$ over a follow-up region $[0, \tau_1] \times [0, \tau_2]$ can be written as

$$U(\beta; \tau_1, \tau_2) = \begin{pmatrix} U_{10}(\beta_{10}; \tau_1) \\ U_{01}(\beta_{01}; \tau_2) \\ U_{11}(\beta_{11}; \tau_1, \tau_2) \end{pmatrix} = 0, \tag{4}$$

where

$$U_{10}(\beta_{10}; t_1) = \sum_{i=1}^{n} \int_0^{t_1} \{X_i(s_1, 0) - \bar{X}(s_1, 0; \beta_{10})\}' N_{1i}(ds_1),$$

$$U_{01}(\beta_{01}; t_2) = \sum_{i=1}^{n} \int_0^{t_2} \{X_i(0, s_2) - \bar{X}(0, s_2; \beta_{01})\}' N_{2i}(ds_2),$$

and

$$U_{11}(\beta_{11}; t_1, t_2) = \sum_{i=1}^{n} \int_0^{t_1} \int_0^{t_2} \{X_i(s_1, s_2) - \bar{X}(s_1, s_2; \beta_{11})\}' N_{1i}(ds_1)N_{2i}(ds_2),$$

and where

$$\bar{X}(s_1, 0; \beta_{10}) = \sum_{i=1}^{n} Y_{1i}(s_1) X_i(s_1, 0) \exp\{X_i(s_1, 0)\beta_{10}\} / \sum_{i=1}^{n} Y_{1i}(s_1) \exp\{X_i(s_1, 0)\beta_{10}\},$$

$$\bar{X}(0, s_2; \beta_{01}) = \sum_{i=1}^{n} Y_{2i}(s_2) X_i(0, s_2) \exp\{X_i(0, s_2)\beta_{01}\} / \sum_{i=1}^{n} Y_{2i}(s_2) \exp\{X_i(0, s_2)\beta_{01}\}, \text{ and}$$

$$\bar{X}(s_1, s_2; \beta_{11}) = \sum_{i=1}^{n} Y_{1i}(s_1) Y_{2i}(s_2) X_i(s_1, s_2) \exp\{X_i(s_1, s_2)\beta_{11}\} / \sum_{i=1}^{n} Y_{1i}(s_1) Y_{2i}(s_2) \exp\{X_i(s_1, s_2)\beta_{11}\}.$$

Note that the solution $\hat{\beta} = (\hat{\beta}'_{10}, \hat{\beta}'_{01}, \hat{\beta}'_{11})'$ to (4) provides an estimator of β under (1)–(3) as derives from the fact that $N_{1i}(ds_1)$ in U_{10} can be replaced by $L_{10i}(ds_1, 0; \beta_{10})$, $N_{2i}(ds_2)$ in U_{01} can be replaced by $L_{10i}(ds_1, 0; \beta_{10})$, $N_{2i}(ds_2)$ and $N_{1i}(ds_1)N_{2i}(ds_2)$ in U_{11} can be replaced by $L_{11i}(ds_1, ds_2; \beta_{11})$ while retaining equality to zero, as follows from some simple algebra. Hence $U(\beta; \tau_1, \tau_2)$ is composed of stochastic integrals of functions of the data with respect to a zero mean process at the 'true' β-value.

The distribution theory for $\hat{\beta}$ is complicated due to the dependence of the 'centering' processes $\bar{X}(\cdot, 0; \beta_{10}), \bar{X}(0, \cdot; \beta_{01})$ and $\bar{X}(\cdot, \cdot; \beta_{11})$ on data from all sampled individuals. However, under independent and identically distributed (i.i.d.) conditions and some regularity conditions these processes converge almost surely to the ratio of the expectations of numerator and denominator terms, denoted here by $\bar{x}(\cdot, 0; \beta_{10}), \bar{x}(0, \cdot; \beta_{01})$ and $\bar{x}(\cdot, \cdot; \beta_{11})$ respectively. In fact, the convergence is at a sufficiently rapid rate as $n \to \infty$ that the centering processes can be replaced by these limits without altering the asymptotic distribution of $n^{-1/2}U(\beta; \tau_1, \tau_2)$. The central limit theorem can therefore be applied to $n^{-1/2}U(\beta; t_1, t_2)$, for all $(t_1, t_2) \in [0, \tau_1] \times [0, \tau_2]$, to show weak convergence to a zero mean Gaussian process with covariance function

$$\Sigma(\beta; t_1, t_2) = E \left(\begin{array}{c} \int_0^{t_1} \{X(s_1, 0) - \bar{x}(s_1, 0; \beta_{10})\}' L_{10}(ds_1, 0; \beta_{10}) \\ \int_0^{t_2} \{X(0, s_2) - \bar{x}(0, s_2; \beta_{01})\}' L_{01}(0, ds_2, 0; \beta_{01}) \\ \int_0^{t_1}\int_0^{t_2} \{X(s_1, s_2) - \bar{x}(s_1, s_2; \beta_{11})\}' L_{11}(ds_1, ds_2; \beta_{11}) \end{array} \right)^{\otimes 2},$$

where E denotes expectation, and $a^{\otimes 2} = aa'$ for column vector a. For these developments (τ_1, τ_2) is required to be in the support of the observed follow-up times $\{(S_{1i}, S_{2i}), i = 1, \ldots, n\}$. A Taylor series expansion of $U(\beta; \tau_1, \tau_2)$ about the 'true' β value then leads under regularity conditions to a mean zero asymptotic Gaussian distribution for $n^{1/2}(\hat{\beta} - \beta)$ with analytic variance estimator $I^{-1}\hat{\Sigma}I^{-1}$ of sandwich form, where I is the product of n^{-1} and the negative

of the matrix of partial derivatives of $U(\beta; \tau_1, \tau_2)$ with respect to β evaluated at $\hat{\beta}$, and $\widehat{\Sigma}$ is an empirical estimator of $\Sigma(\beta; \tau_1, \tau_2)$ that can be written as

$$
\widehat{\Sigma} = n^{-1} \sum_{i=1}^{n} \left(\begin{array}{c} \int_0^{\tau_1} \{X_i(t_1, 0) - \bar{X}(t_1, 0; \hat{\beta}_{10})\}' \hat{L}_{10i}(dt_1, 0; \hat{\beta}_{10}) \\[2mm] \int_0^{\tau_2} \{X_i(0, t_2) - \bar{X}(0, t_2; \hat{\beta}_{01})\}' \hat{L}_{01i}(0, dt_2; \hat{\beta}_{01}) \\[2mm] \int_0^{\tau_1} \int_0^{\tau_2} \{X_i(t_1, t_2) - \bar{X}(t_1, t_2; \hat{\beta}_{11})\}' \hat{L}_{11i}(dt_1, dt_2; \hat{\beta}_{11}) \end{array} \right)^{\otimes 2}, \tag{5}
$$

where $\hat{L}_{10}, \hat{L}_{01}$ and $\hat{L}_{11}$ are obtained from L_{10}, L_{01} and L_{11}, respectively, by evaluating at $\hat{\beta}$ and at empirical estimators of the baseline hazard rates $\Lambda_{10}, \Lambda_{01}$ and Λ_{11} in (1)–(3). Natural empirical estimators $\widehat{\Lambda}_{10}, \widehat{\Lambda}_{01}$ and $\widehat{\Lambda}_{11}$ of these baseline rates have Aalen–Breslow form, and are given by

$$
\widehat{\Lambda}_{10}(t_1, 0) = \int_0^{t_1} \sum_{i=1}^{n} N_{1i}(ds_1) / \sum_{i=1}^{n} Y_{1i}(s_1) \exp\{X_i(s_1, 0)\hat{\beta}_{10}\}, \tag{6}
$$

$$
\widehat{\Lambda}_{01}(0, t_2) = \int_0^{t_2} \sum_{i=1}^{n} N_{2i}(ds_2) / \sum_{i=1}^{n} Y_{2i}(s_2) \exp\{X_i(0, s_2)\hat{\beta}_{01}\}, \text{ and} \tag{7}
$$

$$
\widehat{\Lambda}_{11}(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \sum_{i=1}^{n} N_{1i}(ds_1) N_{2i}(ds_2)
$$
$$
/ \sum_{i=1}^{n} Y_{1i}(s_1) Y_{2i}(s_2) \exp\{X_i(s_1, s_2)\hat{\beta}_{11}\}. \tag{8}
$$

Empirical process methods can be used to show the weak convergence of $n^{1/2}(\widehat{\Lambda}_{10} - \Lambda_{10}), n^{1/2}(\widehat{\Lambda}_{01} - \Lambda_{01}), n^{1/2}(\widehat{\Lambda}_{11} - \Lambda_{11})$ jointly as $n \to \infty$ to a zero mean Gaussian process under (1)–(3). In fact an empirical covariance function estimator for these parameter estimates can be developed. These asymptotic developments follow from modest extensions of the work of Spiekerman and Lin (1998), and Lin et al. (2000). Some detail on these developments and related conditions is given in the Appendix in the more general context of §5.

The marginal single failure hazard rate models (1) and (2) impose constraints on the double failure hazard rate model (3) and visa versa, so that (1) and (2) typically will not be fully consistent with (3) in their respective hazard rate regression components. The solutions to (1)–(3), and the estimators of the baseline rates (6)–(8) may incorporate some asymptotic bias under departure from one or more of the regression models (1)–(3). However, through

the time-varying features of the modeled regression variables in these models, and even further through the use of both time-varying regression variates and time-varying baseline hazard rate stratification, the latter of which can be readily incorporated by replacing (4) by corresponding summations over a fixed number of possibly time-dependent strata, one has the tools to arrange for each of (1), (2) and (3) to provide a suitable fit to available data. Having done so one can expect that estimators of joint survival probabilities and related statistics, for example those that assess the strength of dependency between T_1 and T_2 given Z, to be estimated with little bias. This topic too will be elaborated below. Also, departure from any one, or two, of the hazard rate models (1)–(3) does not adversely affect asymptotic distributional results for estimators of parameters in the remaining hazard rate models.

2.2   Simulation evaluations

Continuous failure times given a single binary covariate z that takes values 0 or 1 with probability 0.5 were generated under the rather specialized Clayton–Oakes regression model (Clayton, 1978; Oakes, 1986) of the form

$$F(t_1, t_2; z) = \{F_0(t_1, 0)^{-\theta} + F_0(0, t_2)^{-\theta} - 1\}^{-e^{z\gamma}/\theta} \qquad (9)$$

for values $\theta\ 0$, where F_0 denotes the survival function at $z = z(0, 0) = 0$. The resulting failure time variates have marginal single failure hazard rates of the form

$$\Lambda_{10}(dt_1, 0; z) = \Lambda_{10}(dt_1, 0)e^{z\gamma} \text{ and } \Lambda_{01}(0, dt_2; z) = \Lambda_{01}(0, dt_2)e^{z\gamma},$$

and double failure hazard rates

$$\Lambda_{11}(dt_1, dt_2; z) = \Lambda_{11}(dt_1, dt_2)e^{z\gamma}(e^{z\gamma} + \theta)/(1 + \theta),$$

so that (1)–(3) is obtained for a binary covariate z with $x(t_1, 0) = x(0, t_2) = x(t_1, t_2) = z$, with $x(t_1, 0) = x(0, t_2) = x(t_1, t_2) = z$, with $\beta_{10} = \beta_{01} = \gamma$, and with $\beta_{11} = \log\{e^{\gamma}(e^{\gamma} + \theta)/(1 + \theta)\}$. Data were generated with unit exponential marginals at $z = 0$, and censoring times that were independent of each other and equally and exponentially distributed with censoring hazard rate, c, chosen to give certain specified uncensored failure fractions for T_1 and T_2, or with no censoring ($c = 0$). Covariate values were generated with probabilities 0.5 for $z = 0$ and $z = 1$.

In implementing (4) τ_1 and τ_2 were specified as the maximal values of S_{1i} and S_{2i} respectively in the sample of size n. Table 1 shows sample means and sample standard deviations for $(\hat{\beta}_{10}, \hat{\beta}_{01}, \hat{\beta}_{11})$ at (θ, γ) values of (2, 0) and (2, log 2) at $n = 250$ or 500 with substantial censoring ($c = 5$), and at $n = 100$ with no censoring, based on 1000 simulations at each configuration. These simulations show little evidence of regression parameter bias, even though there are, for example, only about 13.5 expected double failures at $(\theta, \gamma) = (2, 0)$ and $n = 250$, with substantial censoring ($c = 5$) for each of T_1 and T_2. Also there is good agreement generally between the sample standard deviation for the regression parameter estimates and the average of the standard deviation estimators from the sandwich variance

formula, as well as good proximity to 95% for the associated asymptotic 95% confidence interval coverage. An exception occurs at $(\theta, \gamma) = (2, 0)$ and $n = 250$, where the sample standard deviation for $\hat{\beta}_{11}$ is considerably larger than the average of sandwich standard deviations, presumably reflecting a distribution with heavier tails than the approximating asymptotic normal distribution. The approximation, however, seems adequate at $n = 500$, where the expected number of double failures is about 27. Note that the marginal double failure hazard rate from (9) has a very specialized form, which typically does not agree with the semiparametric model (3) if z is not binary.

Table 2 shows corresponding summary statistics for the cumulative double failure hazard rate estimator $\widehat{\Lambda}_{11}(t_1, t_2; z)$, under (8) and the second Table 1 configuration ($\theta = 2, \gamma = \log 2$), at both $z = 0$ and $z = 1$. One can show the targeted double failure hazard rate to be

$$\Lambda_{11}(t_1, t_2; z) = e^{z\gamma}(e^{z\gamma} + \theta)\{t_1/\theta + t_2/\theta - \log(e^{t_1^{\theta}} + e^{t_2^{\theta}} - 1)/\theta^2\}$$

under the simulation conditions of this subsection. Estimated $\widehat{\Lambda}_{11}$ values are reasonably accurate under the configurations shown, as was also the case at some smaller sample sizes (e.g. $n = 100$ with no censoring). Empirical approximations to asymptotic standard deviation estimates are somewhat low in heavy censoring scenarios, especially close to the coordinate axes or toward distributional tails, and corresponding confidence interval coverage rates tend to be less than nominal levels. These features derive from few preceding double failures close to the axes, and from empty double failure risk sets for some samples toward distributional tails. Hence fairly large sample sizes may be needed for these asymptotic approximations to the distribution of $\widehat{\Lambda}_{11}$ to be accurate.

3   Bivariate Survival Function and Dependency Function Estimation

3.1   Bivariate survival function estimation

Given specifications, such as (1)–(3), for marginal single and double failure hazard rate processes one can define a bivariate process F given Z for all $t_1 \quad 0$ and $t_2 > 0$ by the product integrals

$$F\{t_1, 0; Z(t_1, 0)\} = \prod_0^{t_1} \left[1 - \Lambda_{10}\{ds_1, 0; Z(s_1, 0)\}\right], \text{ and}$$

$$F\{0, t_2; Z(0, t_2)\} = \prod_0^{t_2} \left[1 - \Lambda_{01}\{0, ds_2; Z(0, s_2)\}\right],$$

along the coordinate axes. Away from these axes F given Z is defined by the inhomogeneous Volterra integral equation

$$F\{t_1, t_2; Z(t_1, t_2)\} = F\{t_1, 0; Z(t_1, 0)\} + F\{0, t_2; Z(0, t_2)\} - 1$$

$$+ \int_0^{t_1} \int_0^{t_2} F\{s_1^-, s_2^-; Z(s_1, s_2)\} \Lambda_{11}\{ds_1; ds_2; Z(s_1, s_2)\}, \tag{10}$$

that has a unique Péano series solution given by

$$F\{t_1, t_2; Z(t_1, t_2)\} = \psi\{t_1, t_2; Z(t_1, 0), Z(0, t_2)\}$$

$$+ \sum_{j=1}^{\infty} \int_0^{t_1} \int_{s_{11}}^{t_1} \cdots \int_{s_{1,j-1}}^{t_1} \int_0^{t_2} \int_{s_{21}}^{t_2} \cdots \int_{s_{2,j-1}}^{t_2} \psi\{s_{11}^-, s_{21}^-; Z(s_{11}, s_{21}) \tag{11}$$

$$\} \prod_{m=1}^{j} \Lambda_{11}\{ds_{1m}, ds_{2m}; Z(s_{1m}, s_{2m})\},$$

where $\psi\{t_1, t_2; Z(t_1, 0), Z(0, t_2)\} = F\{t_1, 0; Z(t_1, 0)\} + F\{0, t_2; Z(0, t_2)\} - 1$. Note that F given Z will have a survival function interpretation if Z is composed only of the baseline covariate data, $z(0, 0)$, and evolving covariates that are external to the failure processes (e.g. Kalbfleisch and Prentice, 2002, chapter 6).

Now denote the uncensored T_1 failures in the sample of a size n by $t_{11}, t_{12},..., t_{1I}$, and the uncensored T_2 failures by $t_{21}, t_{22},..., t_{2J}$. The semiparametric model estimators $\widehat{\Lambda}_{10}\{dt_1, 0; Z(t_1, 0)\}, \widehat{\Lambda}_{01}\{0, dt_2; Z(0, t_2)\}$ and $\widehat{\Lambda}_{11}\{dt_1, dt_2; Z(t_1, t_2)\}$ place mass only at the uncensored data grid points (t_{1i}, t_{2j}) within the risk region $[R = \{(t_1, t_2; S_{1\ell} \geq t_1, S_{2\ell} \geq t_2, \text{ for some } \ell \in (1, ..., n)\}]$ of the data, and along the half-lines through the uncensored failure times in either direction beyond the risk region. One can readily estimate F given Z at specified covariate history Z, using a simple recursive procedure as follows: At uncensored failure time grid point $(t_{1i}, t_{2j}) \in R$ one can define

$$\widehat{F}\{\Delta t_{1i}, \Delta t_{2j}; Z(t_{1i}, t_{2j})\} = \widehat{F}\{t_{1i}^-, t_{2j}^-; Z(t_{1i}, t_{2j})\} \widehat{\Lambda}_{11}\{\Delta t_{1i}, \Delta t_{2j}; Z(t_{1i}, t_{2j})\}$$

where $\widehat{\Lambda}_{11}\{\Delta t_{1i}, \Delta t_{2j}; Z(t_{1i}, t_{2j})\} = \widehat{\Lambda}_{11}(\Delta t_{1i}, \Delta t_{2j}) \exp\{X(t_{1i}, t_{2j})\widehat{\beta}_{11}\}$, from which

$$\widehat{F}\{t_{1i}, t_{2j}; Z(t_{1i}, t_{2j})\} = \widehat{F}\{t_{1i}, t_{2j}^-; Z(t_{1i}, t_{2j})\} + \widehat{F}\{t_{1i}^-, t_{2j}; Z(t_{1i}, t_{2j})\}$$
$$- \widehat{F}\{t_{1i}^-, t_{2j}^-; Z(t_{1i}, t_{2j})\} \left[1 - \widehat{\Lambda}_{11}\{\Delta t_{1i}, \Delta t_{2j}; Z(t_{1i}, t_{2j})\} \right]. \tag{12}$$

This expression provides a procedure for calculating $\widehat{F}\{t_{1i}, t_{2j}; Z(t_{1i}, t_{2j})\}$ for a specified covariate history Z, starting with

$$\hat{F}\{t_{1i}, 0; Z(t_{1i}, 0)\} = \prod_{\ell = 1}^{i} [1 - \widehat{\Lambda}_{10}(\Delta t_{1\ell}, 0)e^{X(t_{1\ell}, 0)\hat{\beta}_{10}}], \text{ and}$$

$$\hat{F}\{0, t_{2j}; Z(0, t_{2j})\} = \prod_{m = 1}^{j} [1 - \widehat{\Lambda}_{01}(0, \Delta t_{2m})e^{X(0, t_{2m})\hat{\beta}_{01}}]. \tag{13}$$

Under (1)–(3) $\hat{F}$ given Z will generally provide a strongly consistent estimator of F given Z, and $n^{1/2}(\hat{F} - F)$ given Z will converge as $n \to \infty$ to a zero mean Gaussian process, on the basis of the properties of the univariate Cox model estimators (13) and the continuity and weakly continuous compact differentiability of the Péano series transformation (Gill et al., 1995) from these marginal estimators and $\widehat{\Lambda}_{11}$ to $\hat{F}$, given Z. As such (12) and (13) provide a rather flexible regression generalization, with fixed and external covariates, of a Volterra bivariate survival function estimator that has been attributed to Peter Bickel (Dabrowska, 1988).

3.2 Dependency function estimation

One use of the estimator $\hat{F}$ given Z is for assessing dependency between the two failure time variates given Z. If F given Z has a survival function interpretation one can define, building on the work of Fan et al. (2000), an average cross ratio function estimator $\hat{C}\{t_1, t_2; Z(t_1, t_2)\}$ over $[0, \tau_1] \times [0, \tau_2]$, where (τ_1, τ_2) is in the risk region of the data, by

$$\hat{C}\{t_1, t_2; Z(t_1, t_2)\} = \int_0^{t_1} \int_0^{t_2} \hat{F}\{s_1^-, s_2^-; Z(s_1, s_2)\}\widehat{\Lambda}_{11}\{\Delta s_1, \Delta s_2; Z(s_1, s_2)\}/$$

$$\int_0^{t_1} \int_0^{t_2} \hat{F}\{s_1^-, s_2^-; Z(s_1, s_2)\}\widehat{\Lambda}_{10}\{\Delta s_1, s_2^-; Z(s_1, s_2)\}\widehat{\Lambda}_{01} \tag{14}$$

$$\{s_1^-, \Delta s_2; Z(s_1, s_2)\},$$

which contrasts the double failure rate $\widehat{\Lambda}_{11}$ with the corresponding local independence value $\widehat{\Lambda}_{10}\widehat{\Lambda}_{01}$ given Z, where

$$\widehat{\Lambda}_{10}\{\Delta s_1, s_2^-; Z(s_1, s_2)\} = -\hat{F}\{\Delta s_1, s_2^-; Z(s_1, s_2)\}/\hat{F}\{s_1^-, s_2^-; Z(s_1, s_2)\} \text{ and}$$

$$\widehat{\Lambda}_{01}\{s_1^-, \Delta s_2; Z(s_1, s_2)\} = -\hat{F}\{s_1^-, \Delta s_2; Z(s_1, s_2)\}/\hat{F}\{s_1^-, s_2^-; Z(s_1, s_2)\},$$

with weight function $\hat{F}\{s_1^-, s_2^-; Z(s_1, s_2)\}$ that depends on the failure rates, but not the censoring rates, given Z.

Similarly, one can define, following Oakes (1989) and Fan et al. (2000), an average concordance function estimator between T_1 and T_2 over $[0, \tau_1] \times [0, \tau_2]$ given Z, that takes values in $(-1, 1)$, by

$$\widehat{\mathcal{T}}\{t_1, t_2; Z(t_1, t_2)\}$$

$$= \left[\int_0^{t_1} \int_0^{t_2} \widehat{F}\{s_1^-, s_2^-; Z(s_1, s_2)\}\widehat{F}\{\Delta s_1, \Delta s_2; Z(s_1, s_2)\} \right.$$

$$\left. - \int_0^{t_1} \int_0^{t_2} \widehat{F}\{\Delta s_1, s_2^-; Z(s_1, s_2)\}\widehat{F}\{s_1^-, \Delta s_2; Z(s_1, s_2)\} \right] / \qquad (15)$$

$$\left[\int_0^{t_1} \int_0^{t_2} \widehat{F}\{s_1^-, s_2^-; Z(s_1, s_2)\}\widehat{F}\{\Delta s_1, \Delta s_2; Z(s_1, s_2)\} \right.$$

$$\left. + \int_0^{t_1} \int_0^{t_2} \widehat{F}\{\Delta s_1, s_2^-; Z(s_1, s_2)\}\widehat{F}\{s_1^-, \Delta s_2; Z(s_1, s_2)\} \right].$$

These estimators quite generally inherit strong consistency, and weak Gaussian convergence properties from these same properties for $\widehat{F}$ given Z and the continuity and compact differentiability of the transformations from $\widehat{F}$ to $\widehat{C}$ and from $\widehat{F}$ to $\widehat{\mathcal{T}}$ ..

3.3   Confidence interval and confidence band estimation

The asymptotic properties just stated for survival and dependency function estimators conceptually generate corresponding analytic variance function estimators using the delta function method. However the transformations from marginal single and double hazard rate estimators to the bivariate survival function using (11), and the transformations from the survival function to the average cross ratio and concordance estimators (14) and (15) may be too complex for the delta function approach to be useful. Accordingly we employ a bootstrap resampling approach to estimate confidence intervals and bands for these functions, as well as to estimate confidence bands for marginal single and double failure cumulative hazard functions. The applicability of bootstrap procedures follows from the asymptotic Gaussian properties already cited for regression parameter and baseline hazard function estimators in (10)–(13), and the weakly continuous compact differentiabilty of the Péano series survival function transformation (11) (Gill et al., 1995) and of the transformations (14) and (15) (Fan et al., 2000).

3.4   Simulation evaluation of survival and dependency function estimators

In the special case where all regression parameters in (1)–(3) take value zero, $\widehat{F}$ given Z from (12) and (13) is the previously mentioned Volterra estimator. While nonparametric plug-in estimators of the bivariate survival function due to Dabrowska (1988) and Prentice and Cai (1992) have been shown (Gill et al., 1995) to be nonparametric efficient under the compete independence of (T_1, T_2, C_1, C_2), this property evidently does not hold for the Volterra estimator. On that basis it has been speculated that the Volterra estimator may be 'much inferior' to these other estimators (Gill et al., 1995). To examine this topic further we conducted simulations under (9) with $\gamma = 0$, so that $\beta_{10} = \beta_{01} = \beta_{11} = 0$ with no regression variable influences. As previously T_1 and T_2 were specified as the maximal observed S_{1i} and S_{2i} values respectively in the generated sample.

Table 3 shows summary statistics evaluating the Volterra estimator and comparing it to the Dabrowska estimator, which is also simply calculated recursively using

$$\hat{F}(t_{1i}, t_{2j}) = \frac{\hat{F}(t_{1i}^-, t_{2j})\hat{F}(t_{1i}, t_{2j}^-)}{\hat{F}(t_{1i}^-, t_{2j}^-)} \left\{ \frac{d_{ij}^{00} r_{ij}}{(d_{ij}^{10} + d_{ij}^{00})(d_{ij}^{01} + d_{ij}^{00})} \right\}$$

at all grid points where the denominator components in the factor in curly brackets are positive, and $\hat{F} = 0$ otherwise, again starting with KM marginal survival function estimators. In this expression 'd_{ij}^{10}, d_{ij}^{01} and d_{ij}^{00}' are the numbers of observations known to have '$T_1 = t_{1i}$ and $T_2 > t_{ij}$'; '$T_1 > t_{1i}$ and $T_2 = t_{2j}$'; and '$T_1 > t_{1i}, T_2 > t_{2j}$' respectively, among the r_{ij} individuals at risk at uncensored failure time grid point (t_{1i}, t_{2j}). From Table 3 one can see that both the Volterra and Dabrowska estimators are quite accurate under the specified sampling configurations. The two estimators also appear to have similar corresponding moderate sample efficiencies, even at the complete independence of (T_1, T_2, C_1, C_2), where the Dabrowska estimator is nonparametric efficient. Note that, in contrast to the Dabrowska estimator, the Volterra estimator does not assign negative mass within the risk region of the data. However, it tends to assign more negative mass than does the Dabrowska estimator, to half-lines beyond the risk region. Overall, these simulations provide little basis for choosing between the Volterra and Dabrowska nonparametric estimators of the bivariate survivor function.

Table 4 shows summary statistics for $\hat{F}$ at various follow-up times (t_1, t_2) under (9) and a specific Table 1 configuration $(\theta = 2, \gamma = \log 2)$. The survival function estimators $\hat{F}$ do not show evidence of bias under these simulation conditions, similar to what was observed for smaller sample sizes (e.g., $n = 100$ with no censoring). One could apply the bootstrap procedure to some transformation of $\hat{F}$, such as $\log \hat{F}$, but we applied it directly to $\hat{F}$ in these simulations. Note the good correspondence between sample standard deviation (SD) based on 1000 generated samples at each configuration and the corresponding average of bootstrap SD estimates, based on 200 bootstrap replicates for each generated sample. Also asymptotic 95% confidence interval coverage rates, based on $\hat{F}(t_1, t_2) \pm 1.96$ (bootstrap SD), are close to the nominal levels throughout Table 4.

Supplementary Table 1 compares analytic and bootstrap SD estimators for $\widehat{\Lambda}_{11}$, as well as corresponding 95% confidence interval coverage rates under the same generated samples, and at the same (t_1, t_2) values, as in Table 4. There appears to be good agreement between empirical (sandwich) estimator and bootstrap (200 replicates) standard deviation estimators. Confidence interval coverage rates are low under some configurations, but tend to be a little closer to nominal levels with the bootstrap than with the analytic SD estimators.

Table 5 gives confidence band performance statistics for both $\widehat{\Lambda}_{11}$ and $\hat{F}$, over specified follow-up regions. These were developed by applying a supremum statistic over the confidence region without estimator transformation, for each estimator. Specifically, over a follow-up region $[0, t_1] \times [0, t_2]$ the statistics

$$W_{\Lambda_{11}}(t_1, t_2) = \sup_{[0, t_1] \times [0, t_2]} n^{\frac{1}{2}} |\widehat{\Lambda}_{11}(s_1, s_2) - \Lambda_{11}(s_1, s_2)| \text{ and } W_F(t_1, t_2)$$

$$= \sup_{[0, t_1] \times [0, t_2]} n^{\frac{1}{2}} |\widehat{F}(s_1, s_2) - F(s_1, s_2)|$$

are targeted at specified (t_1, t_2) values. Bootstrap estimates of these quantities are obtained, respectively, by calculating

$$W_{\widehat{\Lambda}_{11}}(t_1, t_2) = \sup_{[0, t_1] \times [0, t_2]} n^{\frac{1}{2}} |\widehat{\Lambda}_{11}^*(s_1, s_2) - \widehat{\Lambda}_{11}(s_1, s_2)| \text{ and } W_{\widehat{F}}(t_1, t_2)$$

$$= \sup_{[0, t_1] \times [0, t_2]} n^{\frac{1}{2}} |\widehat{F}^*(s_1, s_2) - \widehat{F}(s_1, s_2)|$$

where $\widehat{\Lambda}_{11}^*$ and $\widehat{F}^*$ denote bootstrap replicate estimators derived using $\widehat{\Lambda}_{11}$ and $\widehat{F}$. Critical values for an α-level (e.g. $\alpha = 0.95$) confidence region can be estimated as the α percentiles $\widehat{C}_\alpha(\widehat{\Lambda}_{11})$ and $\widehat{C}_\alpha(\widehat{F})$ from the bootstrap replicate supremum statistics $W_{\widehat{\Lambda}_{11}}(t_1, t_2)$ and $W_{\widehat{F}}(t_1, t_2)$ respectively. Corresponding α-level confidence bands are then estimated for region $[0, t_1] \times 0, t_2]$ as

$$\widehat{\Lambda}_{11}(t_1, t_2) \pm n^{-\frac{1}{2}} \widehat{C}_\alpha(\widehat{\Lambda}_{11}) \text{ and } \widehat{F}(t_1, t_2) \pm n^{-\frac{1}{2}} \widehat{C}_\alpha(\widehat{F})$$

respectively.

The simulation summary statistics in Table 5 include bootstrap-based confidence regions for both Λ_{11} and F over certain rectangular follow-up regions, using 200 bootstrap replicates for each generated sample, for each of the latter two configurations of Table 4. Note that the full set of uncensored data grid points for a generated sample was retained for all associated bootstrap samples in the calculation of $W_{\widehat{\Lambda}_{11}}$ and $W_{\widehat{F}}$ statistics. The sample mean and standard deviation of critical value estimates from the 1000 generated samples are shown. Summary statistics for 95% confidence bands for Λ_{11} and F at both $z = 0$ and $z = 1$ are also shown in Table 5. Coverage rates tend to be somewhat low but, considering the size of the standard deviation for the bootstrap critical value estimates, these may improve if a larger number of bootstrap replicates are used.

Supplementary Table 2 provides simulation summary statistics for average cross ratio and average concordance estimators under the same simulation conditions as Table 5. Under these simulation conditions $\widehat{C}(t_1, t_2; z)$ estimates $1 + \theta e^{-z\gamma}$ and $\widehat{\mathcal{T}}(t_1, t_2; z)$ estimates

$\theta e^{-z\gamma}/(\theta e^{-z\gamma} + 2)$ at any $t_1 > 0$ and $t_2 > 0$. As shown in Supplementary Table 2 the cross ratio estimates with follow-up periods $[0, t_1] \times [0, t_2]$ tend to have small upward bias, and average concordance estimators have a small downward bias at these sample sizes, especially under the configuration with substantial censoring. These estimated biases derive in part from moderate sample size distributions that are somewhat skewed, and additional calculations show they can be reduced through simple transformation (e.g., apply asymptotic normal approximation to $\log \hat{C}$, rather than to $\hat{C}$). Bootstrap procedures can again be used to estimate confidence intervals and bands for these dependency function estimators.

4   Composite Outcomes in a Low-fat Dietary Pattern Trial

The Women's Health Initiative (WHI) includes a low-fat dietary pattern randomized controlled trial among 48,835 postmenopausal women (Women's Health Initiative Study Group, 1998). Participating women were in the age range 50–79 at randomization at one of the 40 clinical centers in the US during 1993–1998. Forty percent of the participants were assigned to a low-fat dietary pattern intervention that included goals of reducing dietary fat to 20% of energy, as well as increasing vegetables and fruit to five servings a day and grains to six servings a day. The intervention was administered by nutritionists in groups of size 10–15, with 18 sessions in the first year of the intervention, and quarterly maintenance sessions thereafter over an intervention period that averaged 8.5 years, with subsequent continuing non-intervention follow-up. The other 60% of the participants were assigned to a comparison (control) group, with no dietary intervention. Comparison group women were provided written materials on diet and health only. Breast cancer incidence and colorectal cancer incidence were designated primary outcomes, while coronary heart disease incidence was designated as the secondary trial outcome. Various other clinical outcomes, including mortality from any cause were also ascertained, and used in trial monitoring and reporting.

Chlebowski et al. (2017) recently reported updated analyses of breast cancer incidence (T_1) and total mortality (T_2) from this dietary modification (DM) trial, for both the intervention, and a combined intervention and post-intervention, time periods. Cox models (1)–(3) were applied with $X(t_1, 0) = X(0, t_2) = X, (t_1, t_2) = z$, where z is an indicator for intervention ($z = 1$) or comparison ($z = 0$) randomization assignment, and with baseline stratification on age (5-year intervals) and on randomization status in the companion WHI hormone therapy trials. The (T_1, T_2) failure times were censored by a common value $C_1 = C_2 = C$ equal to the participants follow-up time at the end of the intervention period (3/31/05) or, for a small fraction of women, at the time of earlier loss to follow-up. Since deaths can only follow breast cancer incidence events (i.e. $T_2 \geq T_1$) for a participant, an independent censoring assumption requires specifically that censoring rates for T_2 do not depend on the corresponding T_1 value, an appropriate assumption here since all death ascertainment procedures continued unchanged following a breast cancer diagnosis, including matching to the U.S. National Death Index.

At the end of the intervention period the breast cancer (T_1) hazard ratio (estimated 95% CI) was 0.92 (0.84,1.01), with logrank significance level of $p = 0.09$, with 671 and 1093 incident breast cancer cases in the intervention and comparison groups respectively. The

corresponding values for all-cause mortality (T_2) were 0.98 (0.91, 1.06), with 989 and 1519 deaths in the respective groups. The composite outcome (T_1, T_2) of breast cancer followed by death from any cause had an estimated double failure hazard ratio (95% confidence interval) of 0.64 (0.44, 0.93), with 40 and 94 women experiencing the dual events in the two randomization groups, respectively, during the intervention period.

Note that the composite outcome analysis provides stronger evidence for an intervention benefit (logrank $p = 0.02$) than does the marginal analysis for either outcome separately, in spite of a much smaller number of cases. A corresponding univariate analysis of time from randomization to death attributed to breast cancer has estimated hazard ratio (95% confidence interval) of 0.67 (0.43, 1.06), with logrank $p = 0.08$. There were 27 and 61 deaths attributed to breast cancer in the two groups, respectively, during the intervention period. Another univariate analysis considers death, with classification of whether or not the death followed a breast cancer diagnosis, as a marked point process. This approach leads to a hazard ratio estimate (95% CI) of 0.65 (0.45,0.94), as reported in Chlebowski et al. (2017), which is nearly identical to the double failure hazard ratio estimate given above. In fact the corresponding estimating equations agree except for minor differences in the dual outcome risk set specifications at each death time following breast cancer. The double failure hazard rate model, however, brings potential to address additional questions such as whether the observed intervention influence is primarily through breast cancer incidence or through subsequent survival, and can do so in a manner that retains intention-to-treat interpretation for inferences. For example, suppose that the modeled regression variable in Λ_{11} is extended to $X(t_1, t_2) = \left\{z, z(t_2 - t_1)\right\}$. One then obtains $\hat{\beta}_{11} = (0.226, -0.220)$, with corresponding standard deviation estimates of (0.364, 0.101) from the sandwich-form estimated variance matrix. This gives nominally significant evidence ($p = 0.03$) of a dual outcome hazard ratio that is reduced at larger time periods from breast cancer diagnosis to death. See Chlebowski et al. (2017) for more detailed analyses that also include breast tumor hormone receptor status, subgroup analyses, and longer-term non-intervention follow-up.

For completeness Table 6 shows the estimated survival probability for (T_1, T_2) at follow-up times of three, six, and nine years from randomization for each variate. For this purpose we dropped the baseline hazard rate stratification described above, so that survival function estimators at $z = 0$ and $z = 1$ correspond to the comparison and intervention groups as a whole. Corresponding bootstrap-based 95% confidence intervals and 95% supremum-type confidence bands are also shown, the latter from a rectangular follow-up region with from 0 to 9 years for each failure time variate. These were based on 200 bootstrap replicates with asymptotic approximations applied to F, without transformation. Confidence bands are presented only at follow-up grid points $\{3, 6, 9\} \times \{3, 6, 9\}$ in years. As expected the confidence bands are somewhat wider than corresponding confidence intervals at these follow-up times, especially at short follow-up times. Supplementary Table 3 provides corresponding estimators, bootstrap-based confidence intervals and confidence bands for Λ_{11} using the same bootstrap replicates. Since $T_2 \geq T_1$, these are only of interest on or above the main diagonal.

A second illustration in the same clinical trial illustrates the value of including a focus on marginal hazard rates for T_1 and T_2 beyond counting process intensity modeling. Although diabetes was not a designated outcome in the trial protocol, information on the use of 'pills for diabetes' or 'insulin shots for diabetes' were collected twice annually during the trial intervention period and annually thereafter, through medical update questionnaire. These self-reports were found to be in reasonably good agreement with periodic medication inventories provided by study participants. A total of 45, 579 women were without prevalent diabetes at baseline. Clinical practice dictates the use of diabetes pills as a first line treatment for diabetes, changing to insulin injections if the disease progresses. Cox-type regression models were applied to these data, with baseline rates stratified as described above. An analysis (Howard et al., 2018) of time from randomization to initiation of diabetes pills (T_1) gives a hazard ratio estimate (95% confidence interval) for the low-fat dietary pattern intervention of 0.95 (0.88, 1.02) over the intervention period with $p = 0.13$, with 3179 women developing diabetes. A counting process intensity model was applied to the post-diabetes pills follow-up to ascertain time from randomization to insulin use (T_2). This intensity was modeled to allow a distinct parameter for the intervention hazard ratio, and a baseline hazard rate that retained the original stratification, but also stratified on time-from-randomization to first use of oral diabetes agents (in quartiles). The intervention hazard ratio estimate (95% CI) from this analysis was 0.82 (0.64, 1.04) with a significance level of 0.10 and with 309 women progressing to insulin during the intervention period. This provides some modest evidence that the intervention slowed progression to the more serious type of disease requiring insulin injections, after controlling for time from randomization to the initiation of diabetes pills. A marginal single and double hazard rate analysis of the (T_1, T_2) data was also carried out with the original stratification mentioned above for both time variates and with distinct baseline rates and intervention group regression parameters for the two times. The marginal hazard rate analysis for T_1 is the same as was described above, whereas the marginal hazard rate analysis for time from randomization to diabetes requiring insulin injections (T_2) gave intervention hazard ratio estimate (95% confidence interval) of 0.74 (0.59, 0.94) with intention-to-treat significance level of 0.01. An independent censorship assumption, again with $C_1 \equiv C_2$, is entirely appropriate in this context, so that one obtains considerably stronger evidence of intervention benefit for time from randomization to diabetes requiring insulin than is the case from analysis of either of its component parts; namely, time from randomization to diabetes pills and time from diabetes pills to insulin injections. Moreover, this stronger result arises from a comparison between randomized groups, whereas the time from pills to insulin component of the intensity modeling contrasts groups that may differ in their distributions of time-to-diabetes pills, complicating the associated regression parameter interpretation. The double failure hazard ratio estimate (95% confidence interval) here is nearly identical to that for T_2. Over a longer term follow-up that included a substantial post-intervention period, and a median total follow-up of 17.3 years, the T_1 estimated marginal hazard ratio estimate (95% CI) was 0.96 (0.91, 1.00), while that for T_2 was 0.88 (0.78, 0.99) as was reported in Howard et al. (2018).

5 Higher Dimensional Failure Time Regression Methods

5.1 Hazard rate regression models

With bivariate failure time data there may be natural commonalities in baseline rates and in regression parameters in (1) and (2). For example, in twin studies it may be natural to restrict the baseline hazard rates $\Lambda_{10}(dt_1, 0)$ and $\Lambda_{01}(0, dt_2)$ to be identical, and to require some components of β_{10} and β_{01} to be equal. Following Spiekerman and Lin (1998) we will refer to failure times having a common baseline rate function as failures of the same 'type', and for notational convenience we will redefine the marginal single failure hazard rate regression parameter to have a single value for all failure types by allowing the modeled regression vector to include interaction terms with failure type. Also, we now allow the multivariate failure times to be of arbitrary dimension.

5.2 Regression on marginal single and double failure hazard rates

Suppose that there is an arbitrary number, q, of failure times denoted by $T_1..., T_q$ for each 'study subject,' with a possibly evolving q-dimensional covariate Z. Denote by $z(t_1,...,t_q)$ covariate values at $(t_1,...,t_q)$ and by $Z(t_1,..., t_q) = z(0,..., 0) \vee \{z(s_1,..., s_q); s_1 < t_1,..., s_q < t_q\}$ the covariate history prior to $(t_1,..., t_q)$. Also let M denote a unique mapping from $\{1,..., q\}$ to $\{1,..., K\}$, with $K \quad q$, so that $k = M(j)$ denotes the unique failure type for T_j, out of K possible types, for $j = 1,..., q$. Much of the interest in the study of failure rates on Z typically resides in the marginal single failure hazard rates. Suppose that the single failure hazard rate for T_j at follow-up time t_j, given $Z(0,...,0, t_j, 0,...)$ is modeled by

$$\Gamma_k(dt_j)\exp\{X_k(t_j)\beta\} \tag{16}$$

for $j = 1, \ldots, q$. Note that failures of the same type, k, are assumed to have a common baseline hazard rate function 'Γ_k', which is obtained when the modeled covariate X_k is identically zero, with $X_k(t_j)$ a fixed length row vector which for T_j is formed from $\{t_j; Z(0,...,0, t_j, 0,...0)\}$, and β a corresponding (column) regression vector to be estimated. As noted by Spiekerman and Lin (1998), this parameterization is flexible enough to allow, for example, distinct hazard ratio parameter vectors for each failure type, by including interaction variables with failure type in the specification of X_k.

Similarly suppose that the marginal double failure hazard rate for a pair of failure time variates (T_j, T_h) at follow-up times (t_j, t_h), given $Z(0,...,0, t_j, 0,...,0, t_h, 0,... 0)$, is modeled by

$$\Gamma_{kg}(dt_j, dt_h)\exp\{X_{kg}(t_j, t_h)\gamma\}, \tag{17}$$

for each $1 \quad j < h \quad q$, where $k = M(j)$ and $g = M(h)$ are the failure types for T_j and T_h, respectively. In (17) $X_{kg}(t_j, t_h)$ is a fixed length row regression vector which for (T_j, T_k) is formed from $\{t_j, t_h; Z(0,...,0, t_j, 0,...,0, t_h, 0,...,)\}$, γ is a corresponding column double failure hazard ratio parameter to be estimated, and 'Γ_{kg}' is a baseline double failure hazard rate function that is obtained at $X_{kg} \equiv 0$.

In this formulation the failure times $T_1, ..., T_q$ can occur along the same or different failure time axes, but failures of the same type are required to fall on the same time axis. For the parameters in (16) and (17) to have a useful interpretation an independent censorship condition, given Z, needs to be met. Hence we assume that lack of censoring in $[0, t_j)$ can be added to the single failure hazard rate conditioning without affecting (16) for any $j = 1, ..., q$, and lack of censoring in $[0, t_j) \times [0, t_h)$ can be added to the double failure hazard rate conditioning without affecting (17), for any (t_j, t_h) and $1 \leq j < h \leq q$.

Now consider a random sample $\{(S_{ji}, \delta_{ji}), j = 1, ..., q; Z(S_{1i}, ..., S_{qi})\}$, for $i = 1, ..., n$ from a study cohort, where $S_{ji} = T_{ji} \wedge C_{ji}$ is the minimum of the jth failure time T_{ji} and a corresponding potential censoring time C_{ji} for the ith individual, and $\delta_{ji} = I[S_{ji} = T_{ji}]$. From these one can define counting processes N_{ji} and 'at risk' processes Y_{ji} by

$$N_{ji}(dt_j) = \begin{cases} 1 & \text{if } S_{ji} = t_j \text{ and } \delta_{ji} = 1 \\ 0 & \text{otherwise} \end{cases} ; Y_{ji}(t_j) = \begin{cases} 1 & \text{if } S_{ji} \geq t_j \\ 0 & \text{otherwise} \end{cases}$$

for $j = 1, ..., q$ and $i = 1, ..., n$. Missing failure times can be accommodated by setting the pertinent C_{ji} value equal to zero.

Similar to Spiekerman and Lin (1998) one can define an estimating equation for the marginal single failure hazard ratio parameter β by

$$\sum_{i=1}^{n} \sum_{j=1}^{q} \sum_{k=1}^{K} I\{M(j) = k\} \int_0^{\tau_k} \{X_{ki}(t_j) - \bar{X}_k(t_j; \beta)\}' N_{ji}(dt_j) = 0. \tag{18}$$

Also, a corresponding estimating equation for the double failure hazard rates parameter γ can be written

$$\sum_{i=1}^{n} \sum_{j=1}^{q} \sum_{h=j+1}^{q} \sum_{k=1}^{K} \sum_{g=k+1}^{K} I\left\{M(j) = k\right\} I\left\{M(h) = g\right\} \tag{19}$$
$$\int_0^{\tau_k} \int_0^{\tau_g} \{X_{kgi}(t_j, t_h) - \bar{X}_{kg}(t_j, t_h; \gamma)\}' N_{ji}(dt_j) N_{hi}(dt_h) = 0.$$

In these expressions (τ_k, τ_g) are such that
$P\{S_{ji} \geq \tau_k, M(j) = k, S_{hi} \geq \tau_g, M(h) = g; Z(0, ..., t_j, 0, ..., 0, t_h, 0, ...0)\} > 0$ for some (j, h), for each $1 \leq k \leq g \leq K$ for theoretical developments, but each τ_k can evidently be taken to be the maximal observed S_{ji} value, where $k = M(j)$, in application. Also the 'centering' variates in (18) are

$$\bar{X}_k(t_j; \beta) = Q_k^{(1)}(t_j; \beta) / Q_k^{(0)}(t_j; \beta),$$

where $Q_k^{(\ell)}(t_j; \beta) = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{q} I\{M(j) = k\} Y_{ji}(t_j) X_{ki}(t_j)^{(\ell)} \exp\{X_{ki}(t_j)\beta\}$, for $\ell = 0, 1, 2$ with $a^{(0)} = 1$, $a^{(1)} = a$ and $a^{(2)} = a'a$ a row vector a, while those in (19) are

$$\bar{X}_{kg}(t_j, t_h; \gamma) = Q_{kg}^{(1)}(t_j, t_h; \gamma) / Q_{kg}^{(0)}(t_j, t_h; \gamma),$$

where $Q_{kg}^{(\ell)}(t_j, t_h; \gamma) = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{q} \sum_{h=j+1}^{q} I\{M(j) = k\} I\{M(h) = g\}$
$\left\{Y_{ji}(t_j) Y_{hi}(t_h) X_{kg}(t_j, t_h)^{(\ell)} \exp\{X_{kg}(t_j, t_h)\gamma\}\right\}$, for $\ell = 0, 1, 2$.

The utility of (18) and (19) as estimating functions derives from the fact that each $N_{ji}(dt_j)$ in (18) can be replaced by $L_{ji}(dt_j; \beta)$ where

$$L_{ji}(t_j; \beta) = \int_0^{t_j} \left[N_{ji}(ds_j) - \sum_{k=1}^{K} I\{M(j) = k\} Y_{ji}(s_j) \exp\{X_{ki}(s_j)\beta\} \Gamma_k(ds_j) \right]$$

while retaining equality to zero under (16), and similarly each $N_{ji}(dt_j) N_{hi}(dt_h)$ in (19) can be replaced by $L_{jhi}(dt_j, dt_h; \gamma)$ where

$$L_{jhi}(t_j, t_h; \gamma) = \int_0^{t_j} \int_0^{t_h} \left[N_{ji}(ds_j) N_{hi}(ds_h) - \sum_{k=1}^{K} \sum_{g=k+1}^{K} I\{M(j) = k\} I\{M(h) = g\} Y_{ji}(s_j) Y_{hi}(s_h) \right.$$
$$\left. \exp\{X_{kgi}(s_j, s_h)\gamma\} \Gamma_{kg}(ds_j, ds_h) \right],$$

while retaining equality to zero under (17). It follows that the product of $n^{-1/2}$ and left sides of (18) and (19) are stochastic integrals of sample variates with respect to processes L_{ji} and L_{jhi} that have zero means under (16) and (17) respectively at the true (β, γ) values. Moreover, it turns out that under i.i.d. conditions for the processes $\{(N_{ji}, Y_{ji}), j = 1, \ldots, q, Z_i(S_{ji}, \ldots, S_{jq})\}$ for $i = 1, \ldots, n$ that the centering variates in (18) and (19) can be replaced by their almost sure limits.

$$\bar{x}_k(t_j; \beta) = q_k^{(1)}(t_j; \beta) / q_k^{(0)}(t_j; \beta) \text{ and } \bar{x}_{kg}(t_j, t_h; \gamma) = q_{kg}^{(1)}(t_j, t_h; \gamma) / q_{kg}^{(0)}(t_j, t_h; \gamma),$$

where $q_k^{(\ell)}$ and $q_{kg}^{(\ell)}$ are expectations of $Q_k^{(\ell)}$ and $Q_{kg}^{(\ell)}$ respectively for $\ell = 0, 1, 2$, without altering the asymptotic distribution of the left sides of (18) and (19). It then follows further that the left sides of (18) and (19) behave, for large n, like a sum of i.i.d. variates to which the central limit theorem applies under modest additional regularity conditions. From this $n^{-1/2}$ times these left sides converges to a zero mean Gaussian variate at the 'true' values for (β, γ) under (16) and (17). The variance matrix for this Gaussian variate quite generally can be consistently estimated by

$$
\widehat{\Sigma} = n^{-1} \sum_{i=1}^{n} \left(\begin{array}{c} \sum_{j=1}^{q} \sum_{k=1}^{K} I\left\{M(j) = k\right\} \int_{0}^{\tau_k} \{X_{ki}(t_j) - \bar{X}_k(t_j; \hat{\beta})\}' \hat{L}_{ji}(dt_j; \hat{\beta}) \\ \sum_{j=1}^{q} \sum_{h=j+1}^{q} \sum_{k=1}^{K} \sum_{g=k+1}^{K} I\left\{M(j) = k\right\} I\left\{M(h) = g\right\} \\ \int_{0}^{\tau_k} \int_{0}^{\tau_g} \{X_{kgi}(t_j, t_h) - \bar{X}_{kg}(t_j, t_h; \hat{\gamma})\}' \hat{L}_{jhi}(dt_j, dt_h; \hat{\gamma}) \end{array} \right)^{\otimes 2},
$$

where L_{ji} and L_{jhi} denote L_{ji} and L_{jhi} respectively evaluated at $(\hat{\beta}, \hat{\gamma})$ and at Aalen–Breslow estimators of baseline hazard functions given by

$$
\widehat{\Gamma}_k(t_k; \hat{\beta}) = \sum_{i=1}^{n} \sum_{j=1}^{q} I\left\{M(j) = k\right\} \int_{0}^{t_k} N_{ji}(ds_k) / \left\{n Q_k^{(0)}(s_k; \hat{\beta})\right\}, \text{ and}
$$

$$
\widehat{\Gamma}_{kg}(t_k, t_g; \hat{\gamma}) = \sum_{i=1}^{n} \sum_{j=1}^{q} \sum_{h=j+1}^{q} \int_{0}^{t_k} \int_{0}^{t_g}
$$
$$
I\left\{M(j) = k\right\} I\left\{M(h) = g\right\} N_{ji}(ds_k) N_{hi}(ds_g) / \left\{n Q_{kg}^{(0)}(s_k, s_g; \hat{\gamma})\right\},
$$

for $k = 1, \ldots, K$ and for $1 \leq k \leq g \leq K$, respectively.

Taylor series expansions of the left sides of (18) and (19) about the true (β, γ) values then lead, under regularity conditions, to a zero mean asymptotic normal distribution for

$$
n^{\frac{1}{2}} \left(\begin{array}{c} \hat{\beta} - \beta \\ \hat{\gamma} - \gamma \end{array} \right)
$$

with variance matrix consistently estimated by $\hat{A}^{-1} \hat{\Sigma} \hat{A}^{-1}$ where $\hat{A}$ is the product of n^{-1} and the negative of the derivative matrix of the left sides of (18) and (19) with respect to (β, γ). Specifically $n\hat{A}$ is a block diagonal matrix with entries

$$
\sum_{k=1}^{K} \int_{0}^{\tau_k} \left[Q^{(2)}(t_k; \hat{\beta}) / Q^{(0)}(t_k, \hat{\beta}) - \{Q^{(1)}(t_k; \hat{\beta}) / Q^{(0)}(t_k; \hat{\beta})\}^{\otimes 2} \right] \sum_{j=1}^{q} I\left\{M(j) = k\right\} \sum_{i=1}^{n} N_{ji}(dt_k)
$$

in the upper left, with entries

$$
\sum_{k=1}^{K} \sum_{g=k+1}^{K} \int_{0}^{\tau_k} \int_{0}^{\tau_g} \left[Q^{(2)}(t_k, t_g; \hat{\gamma}) / Q^{(0)}(t_k, t_g; \hat{\gamma}) - \{Q^{(1)}(t_k, t_g; \hat{\gamma}) / Q^{(0)}(t_k, t_g; \gamma)\}^{\otimes 2} \right]
$$
$$
\sum_{j=1}^{q} \sum_{h=j+1}^{q} I\left\{M(j) = k\right\} I\left\{M(h) = g\right\} \sum_{i=1}^{n} N_{ji}(dt_k) N_{hi}(dt_g)
$$

in the lower right, and with zero matrices in the off-diagonal blocks, so that one can write

$$\hat{A} = \begin{pmatrix} \hat{A}_1 & 0 \\ 0 & \hat{A}_2 \end{pmatrix}.$$

Empirical process methods can also be used to show $n^{1/2}(\widehat{\Gamma}_k - \Gamma_k)$, for $k = 1, \dots, K$ and $n^{1/2}(\widehat{\Gamma}_{kg} - \Gamma_{kg})$ for $1 \le k < g \le K$ to converge jointly to a zero mean Gaussian processes under (16) and (17), and a sandwich-type covariance process estimator can be specified for this set of parameter estimates. These asymptotic developments again follow from modest extensions of Spiekerman and Lin (1998) and Lin et al. (2000). Some related detail is given in the Appendix. As in the previous section, bootstrap resampling procedures can be used for supremum-type confidence band estimation for marginal single and double cumulative hazard rates, or for confidence intervals or bands for bivariate survival function estimators, and for pairwise cross ratio or concordance functions, given Z, for any $1 \quad k \quad g \quad K$.

It can also be remarked that these asymptotic results assume the marginal single and double failure rate models (16) and (17) to hold simultaneously. Note however, that the asymptotic properties for $\hat{\beta}$ and $\widehat{\Gamma}_k$, for $k = 1, \dots, K$ hold under (16) even under departure from (17), and those for $\hat{\gamma}$ and $\widehat{\Gamma}_{kg}$, for all $1 \quad k \quad g \quad K$ hold under (17) even under departure from (16), providing some flexibility in the modeling and interpretation of the respective single and double failure hazard rates. For example, the marginal single failure hazard ratio factor $\exp\{X_k(t)\hat{\beta}\}$ may have an interpretation as an average failure type k hazard ratio for the modeled covariate even if (16) is oversimplified and (17) fails to hold, and similarly for $\exp\{X_{kg}(t_1, t_2)\hat{\gamma}\}$ under an oversimplified double failure hazard rate model (17) and departure from (16). However, when the fitted marginal single and double failure hazard rates are brought together to estimate bivariate survival functions and pairwise dependency functions given Z, some care may be needed to ensure an adequate fit of (16) and (17) to available data, as will be considered further in Section 6 below.

Note also that mixed continuous and discrete failure times are included in the methodology described above, subject to the models (16) and (17), and the sandwich-type variance estimators and other weak convergence results mentioned above adapt appropriately to the nature of the failure time variates.

It may be possible to improve the efficiency of β estimators by introducing weights into the left side of (18) (e.g., Lin et al., 2000) possibly using the fitted marginal double failure hazard rates as a source of weighting information. Efficiency gains are likely to be small however, unless dependencies among the failure times are strong and censoring is not too severe. Related asymptotic results extend in a straightforward manner under some additional regularity conditions, following Lin et al. (2000), provided any dependence of the weights on β in estimating equation (18) is fixed at the parameter estimate $\hat{\beta}$ described above. Further study of the preferred form of weights that could be included in (18) and of their value for enhancing estimator efficiency, would be worthwhile.

6 Additional Aspects of Hazard Ratio Regression Parameter Modeling and Estimation

6.1 Model misspecification

Some readers may find it problematic that the single and double failure hazard rate models (16) and (17) may be mutually incompatible in that there may be no proper 'survival' function F given Z for which these models are simultaneously obtained. This issue arises also for mean and covariance parameter estimation using estimating equations with uncensored outcomes (e.g. Liang and Zeger, 1986). If this situation arises then one or both of (16) and (17) are misspecified, and, as usual, one can then expect some bias in estimators of related parameters, such as F given Z. An advantage of the semiparametric models (16) and (17), however, is that the unspecified baseline hazard rate functions provide valuable flexibility to these models, with restrictions entering only through the parametric form of the hazard ratio factors. The time-dependent covariate option allows the data analyst to adapt these hazard ratio factors to available data, and time-dependent baseline hazard rate stratification options allow even more flexible modeling. Hence, under careful modeling one can expect to obtain estimated single and double failure hazard rate estimators that are consistent with available data. These estimators uniquely determine estimators $\hat{F}$ given Z for all univariate and bivariate failure times, and these too will then be consistent with available data.

From a practical point of view a data analyst is likely to just include some simple time-dependent terms in the modeled single and double failure regression vectors in (16) and (17). We considered a generalization of the bivariate survival function model (9) where the single failure hazard rates for the binary covariate z are correctly modeled, but the double failure hazard rate is not, to examine the bias associated with this model misspecification, and to examine the extent to which it can be mitigated by the inclusion of the simple time-dependent components $z \log t_1$ and $z \log t_2$ in the respective regression vectors $X(t_1, 0)$ and $X(0, t_2)$, and the inclusion of both of these time-dependent terms in $X(t_1, t_2)$, in the models (1)–(3).

The joint survival function considered was the Clayton-Oakes model

$$F(t_1, t_2, z) = \left[F_0(t_1, 0)^{-\theta \exp\{z(\beta_{10} - \gamma)\}} + F_0(0, t_2)^{-\theta \exp\{z(\beta_{01} - \gamma)\}} - 1 \right]^{-e^{z\gamma}/\theta} \tag{20}$$

for $\theta \geq 0$, with F_0 again denoting the survival function at $z = 0$. This class of models has the same single failure hazard rates as (9), and the same cross ratio function $1 + \theta e^{-z\gamma}$, but the double failure hazard rate model has the more complex form

$$\Lambda_{11}(dt_1, dt_2; z) = \Lambda_{11}(dt_1, dt_2)e^{z\gamma}(e^{z\gamma} + \theta)/(1 + \theta)$$

$$\left|\frac{F_0(t_1, 0)^{-\theta} + F_0(0, t_2)^{-\theta} - 1}{F_0(t_1, 0)^{-\theta\exp\{z(\beta_{10} - \gamma)\}} + F_0(0, t_2)^{-\theta\exp\{z(\beta_{01} - \gamma)\}} - 1}\right|^2$$

$$F_0(t_1, 0)^{-\theta[\exp\{z(\beta_{10} - \gamma)\} - 1]} F_0(0, t_2)^{-\theta[\exp\{z(\beta_{01} - \gamma)\} - 1]},$$

which departs from (3) under departure from $\beta_{10} = \beta_{01} = \gamma$. Table 7 shows some simulation results for estimating F given Z, at both $z = 0$ and $z = 1$ for $\beta_{10} = 0, \beta_{01} = \log2, \theta = 2$, and $\gamma = \log2$, either with $x(t_1, 0) = x(0, t_2) = x(t_1, t_2) = z$ as before, or with $x(t_1, 0) = (z, z\log t_1), x(0, t_2) = (z, z\log t_2)$ and $x(t_1, t_2) = (z, z\log t_1, z\log t_2)$. From the left side of Table 7 one sees that the biases in $\hat{F}$ given z are minimal in the heavy censoring scenario, even without time-dependent regression variables, whereas bias is evident away from the origin in the uncensored data scenario where the model misspecification has more influence in the tails of the survival function. Much of this bias is avoided by the inclusion of these simple time-varying regression variable that allow the single and double failure hazard ratio for $z = 1$ versus $z = 0$ to be power functions of t_1 and t_2. Note that the sample standard deviations for $\hat{F}$ given z are little affected by the inclusion of these time-dependent variables. Corresponding estimators of average cross ratios and average concordances incorporate somewhat greater biases under these sampling configurations, but these biases too were considerably reduced by the inclusion of the time-dependent components of the modeled regression variables. Results were similar at various other parameter values, sample sizes and censoring configurations. Time-dependent regression variable zt_1 and zt_2, instead of $z \log t_1$ and $z \log t_2$ were also considered, with very similar bias reduction properties for these choices, under the simulation model (20).

An additional simulation was conducted under (20) and the same parameter configuration described above, but with regression vectors augmented to include a standard normal variate in addition to the binary regression variable, with the two regression variables having identical parameter values β_{10}, β_{01} and γ. Analyses that included modeled regression variables $z \log t_1$ and $z \log t_2$ for marginal single failure hazard rates, and $z \log(t_1 + t_2 + 1)$ for the marginal double failure hazard rate demonstrated good agreement between sample standard deviations and the average of standard deviation estimators from sandwich estimation–based standard deviation estimators, and good agreement of sandwich estimation–based 95% confidence intervals with nominal levels for targeted parameters, based on 1000 simulated data sets.

6.2 Higher dimensional marginal hazard rate regression estimation

Marginal hazard rate regression models analogous to (16) and (17) can also be considered for trivariate and higher dimensional marginal hazard rates. The methods of the preceding section generalize naturally to the estimation of hazard ratio regression parameters and baseline hazard rates for subsets of the failure times $(T_1, ..., T_q)$ for any $q \geq 1$. Moreover, the survival function F given Z for $(T_1, ..., T_q)$ at a specified q-dimensional covariate history, with fixed or external covariates, can be readily estimated in a recursive fashion. For example, one can write

$$F\{t_1, ..., t_q; Z(t_1, ..., t_q)\} = \psi_Z(t_1, ..., t_q) +$$
$$\int_0^{t_1} \cdots \int_0^{t_q} F\{s_1^-, ..., s_q^-; Z(s_1, ..., s_q)\}\Lambda_{1...1}\{ds_1, ..., ds_q; Z(s_1, ..., s_q)\}$$

where $\psi_Z(t_1, ..., t_q)$ depends only on marginal distributions of F given Z of dimension less than q. This inhomogenous Volterra integral equation has a unique solution as a function of

$$\psi_Z(t_1, ..., t_q) = (-1)^{q-1} + (-1)^{q-2} \sum_{i=1}^{q} F\{0, ..., t_i, 0, ..., 0; Z(0, ..., t_i, 0, ..., 0)\}$$

$$+ (-1)^{q-3} \sum_{i=1}^{q} \sum_{j=1}^{q} F\{0, ..., t_i, 0, ..., 0, t_j, 0, ..., 0; Z(0, ..., 0, t_i, 0, ..., t_j, 0, ..., 0)\}$$

$$+ \cdots + (-1)^{0} \sum_{\substack{i_1=1 \\ i_1 < i_2 < \cdots < i_{q-1}}}^{q} \cdots \sum_{i_{q-1}=1}^{q} F\{0, ..., 0, t_{i_1}, 0, ..., 0, t_{i_{q-1}}, 0, ..., 0;$$

$$Z\left(0, ..., 0, t_{i_1}, 0..., 0, t_{i_{q-1}}, 0..., 0\right)\}$$

and the q-variate hazard rate regression model $\Lambda_{1...1}$ in Péano series form, leading to strongly consistent and weakly Gaussian convergent estimators of F given Z by plugging in marginal hazard rate regression estimators for hazard rates of all dimensions up to q, starting with Cox model marginal single failure hazard rate estimators.

Note, however, that marginal q-variates hazard rate estimators have precision that depends directly on the number of individuals experiencing a q-variate failure. In many applications, for example in epidemiologic cohort studies with failure times constituting q specific clinical outcomes, the data for estimating high-dimensional marginal hazard rates will be too sparse to be useable. In fact, the most useful and interpretable regression information will often derive from marginal single and double failure rate estimation, analogous to mean and

covariance parameter estimation in uncensored data regression settings (e.g. Liang and Zeger, 1986; Prentice and Zhao, 1991).

6.3 Summary and Concluding remarks

In summary, the methods provided here aim to fill an important gap in the various possible extensions of the univariate failure time Cox model to multivariate failure time data. The proposed marginal methods are based on semiparametric multiplicative form regression models for marginal single and double, and potentially higher order, failure hazard rates, where marginal implies that possibly-evolving covariate histories are included in the hazard rate conditioning, but the evolving failure time counting process for the 'individual' (correlated set of measurements) is not included. These methods, along with models of a similar form for the counting process intensity, which does condition on the preceding counting process history, provide flexible tools for the analysis of multivariate failure time regression data. The present marginal methods allow separate censoring processes to apply to the components of the multivariate failure time variable, and allow failure time components of different types to fall on unrelated time axes, provisions that are not available for martingale-based distribution theory for counting process intensity models. On the other hand, intensity process modeling allows censoring rates to depend on the prior counting process data for the correlated set, while somewhat stronger censoring requirements apply to the marginal hazard rate methods considered here.

The applicability of these stronger censoring requirements can be examined by applying models of the form (16) and (17) to marginal single and double failure censoring rates, while extending the conditioning event to include aspects of the preceding failure counting process for the 'individual' in addition to the preceding covariate history. A dependence of these censoring rates on the prior counting process history would suggest departure from independent censoring given Z.

The marginal methods can also be viewed as extending copula model methods to include a semiparametric class of dependency models, including models that can depend on an evolving covariate process. Additionally, the proposed marginal methods build upon the marginal single failure regression methods of Lin, Wei, and colleagues, while including higher dimensional marginal hazard rate regression models, and do so using straightforward computations that extend those used in this earlier work, and in Cox's (1972) seminal paper. As a byproduct, these methods yield semiparametric bivariate survival function estimators, and related cross ratio and concordance dependency function estimators, with fixed or external covariates, that are considerably more flexible than corresponding estimators previously available using copula and frailty regression model approaches. Furthermore the relationship of marginal double failure hazard rates to covariates will often be readily interpretable, and may lead to novel insights; for example, into intervention effects and related intervention mechanisms in a clinical trial context.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Biography

Ross L. Prentice is Professor, Fred Hutchinson Cancer Research Center, Seattle WA 98109 (rprentic@whi.org) and Shanshan Zhao is a Principal Investigator, National Institute of Environmental Health Sciences, Research Triangle Park NC 27709 (shanshan.zhao@nih.gov).

References

Aalen O, Borgan 0, and Gjessing H. (2010). Survival and Event History Analysis. Springer Science & Business Media.

Andersen PK, Borgan 0, Gill RD, and Keiding N. (1993). Statistical Models Based on Counting Processes. New York: Springer-Verlag.

Andersen PK and Gill RD (1982). Cox's regression model for counting processes: a large sample study. The Annals of Statistics 10(4), 1100–1120.

Bandeen-Roche K. and Ning J. (2008). Nonparametric estimation of bivariate failure time associations in the presence of a competing risk. Biometrika 95 (1), 221–232. [PubMed: 20305739]

Chlebowski RT, Aragaki AK, Anderson GL, Thomson CA, Manson JE, Simon MS, Howard BV, Rohan TE, Snetselar L, Lane D, Barrington W, Vitolins MZ, Womack C, Qi L, Hou L, Thomas F, and Prentice RL (2017). Low-fat dietary pattern and breast cancer mortality in the women's health initiative randomized controlled trial. Journal of Clinical Oncology 35(25), 2919–2926. [PubMed: 28654363]

Clayton DG (1978). Model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65, 141–151.

Cox DR (1972). Regression models and life-tables (with discussion). Journal of the Royal Statistical Society. Series B (Methodological) 34 (2), 187–220.

Dabrowska DM (1988). Kaplan–Meier estimate on the plane. Annals of Statistics 16, 1475–1489.

Duchateau L. and Janssen P. (2010). The Frailty Model. New York: Springer-Verlag.

Fan J, Prentice RL, and Hsu L. (2000). A class of weighted dependence measures for bivariate failure time data. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 62(1), 181–190.

Gill RD, van der Laan MJ, and Wellner JA (1995). Inefficient estimators of the bivariate survival function for three models. Ann. Inst. H. Poincare Probab. Statist 31 (3), 545–597.

Hougaard P. (2000). Analysis of Multivariate Survival Data, Volume 564. Springer New York.

Howard BV, Aragaki AK, Tinker LF, Allison M, Hingle MD, Johnson KC, Manson JE, Shadyab AH, Shikany JM, Snetselaar LG, Thomson CA, Zaslavsky O, and Prentice RL (2018). A low-fat dietary pattern and diabetes: A secondary analysis from the Women's Health Initiative Dietary Modification Trial. Diabetes Care 41, 680–687. [PubMed: 29282203]

Hu T, Nan B, Lin X, and Robins JM (2011). Time-dependent cross ratio estimation for bivariate failure times. Biometrika 98(2), 341–354. [PubMed: 22822258]

Kalbfleisch JD and Prentice RL (2002). The Statistical Analysis of Failure Time Data, Second Edition. New York: Wiley and Sons.

Liang K-Y and Zeger SL (1986). Longitudinal data analysis using generalized linear models. Biometrika 73(1), 13–22.

Lin D, Wei L, Yang I, and Ying Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 62(4), 711–730.

Nan B, Lin X, Lisabeth LD, and Harlow SD (2006). Piecewise constant cross-ratio estimation for association of age at a marker event and age at menopause. Journal of the American Statistical Association 101 (473), 65–77.

Nelsen RB (2007). An Introduction to Copulas (Seconded.). New York: Springer-Verlag.

Oakes D. (1986). Semiparametric inference in a model for association in bivariate survival data. Biometrika 73 (2), 353–361.

Oakes D. (1989). Bivariate survival models induced by frailties. Journal of the American Statistical Association 84 (406), 487–493.

Prentice RL and Cai J. (1992). Covariance and survivor function estimation using censored multivariate failure time data. Biometrika 79(3), 495–512.

Prentice RL and Zhao LP (1991). Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. Biometrics, 825–839. [PubMed: 1742441]

Shih JH and Louis TA (1995). Inferences on the association parameter in copula models for bivariate survival data. Biometrics, 1384–1399. [PubMed: 8589230]

Spiekerman CF and Lin D. (1998). Marginal regression models for multivariate failure time data. Journal of the American Statistical Association 93(443), 1164–1175.

Wienke A. (2011). Frailty Models in Survival Analysis. Boca Raton: Chapman and Hall/CRC Press.

Women's Health Initiative Study Group (1998). Design of the Women's Health Initiative clinical trial and observational study. Controlled Clinical Trials 19(1), 61–109. [PubMed: 9492970]

Table 1:

Simulation summary statistics for marginal single failure hazard ratio parameters β_{10} and β_{01} and for double failure hazard ratio parameter β_{11}, and for related estimators from sandwich-form variance estimators, under the Clayton–Oakes model (9) with $\beta_{10} = \beta_{01} = \gamma$ and a binary covariate z. This model implies marginal hazard ratio (for $z = 1$ versus $z = 0$) of e^γ for both T_1 and T_2, and double failure hazard ratio of $e^\gamma(e^\gamma + \theta)/(1 + \theta)$.

$(\theta = 2, \gamma = 0)$

Sample size (n)		250				500				100			
T_1 and T_2 failing %		16.7				16.7				16.7			
Double failure %		5.3				5.3				5.3			
	True	Sample Mean[a]	Sample SD[b]	Sandwich[b] SD	95% CI[b] Coverage	Sample Mean	Sample SD	Sandwich SD	95% CI Coverage	Sample Mean[a]	Sample SD	Sandwich SD	95% CI Coverage
β_{10}	0	0.003	0.315	0.316	0.953	0.006	0.227	0.222	0.951	−0.001	0.213	0.203	0.933
β_{01}	0	0.007	0.327	0.317	0.948	−0.008	0.224	0.221	0.952	−0.005	0.216	0.203	0.943
β_{11}	0	−0.001	1.222	0.606	0.961	−0.007	0.440	0.420	0.949	−0.003	0.282	0.263	0.928

$(\theta = 2, \gamma = \log 2)$

Sample size (n)		250				500				100			
T_1 and T_2 failing %		22.6				22.6				22.6			
Double failure %		8.2				8.2				8.2			
	True	Sample Mean	Sample SD	Sandwich SD	95% CI Coverage	Sample Mean	Sample SD	Sandwich SD	95% CI Coverage	Sample Mean	Sample SD	Sandwich SD	95% CI Coverage
β_{10}	0.693	0.706	0.281	0.281	0.966	0.705	0.203	0.198	0.941	0.696	0.222	0.212	0.942
β_{01}	0.693	0.707	0.287	0.282	0.959	0.697	0.199	0.197	0.955	0.696	0.223	0.212	0.937
β_{11}	0.981	1.056	0.798	0.525	0.956	1.004	0.381	0.364	0.951	0.986	0.315	0.288	0.929

[a] Based on 1000 simulations at each sampling configuration.

[b] Abbreviations: Sample SD is sample standard deviation of the regression parameter estimates; Sandwich SD is the corresponding average of standard deviation estimates derived from the sandwich form estimator of variance for $\hat{\beta}$, and 95% CI coverage is the fraction of simulated samples from which the approximate 95% confidence interval, formed from $\hat{\beta}$ and its sandwich-form variance estimator, includes the true β.

Table 2:

Simulation summary statistics* for double failure hazard estimator $\widehat{\Lambda}_{11}$ at selected marginal survival probabilities given z, under the Clayton-Oakes model (9) with $\beta_{10} = \beta_{01} = \gamma = \log 2$ and dependency parameter $\theta = 2$, at both $z = 0$ and $z = 1$.

Sample size (n)		500				1000				250			
T_1 and T_2 Failing %		22.6				22.6				100			
		$\widehat{\Lambda}_{11}$				$\widehat{\Lambda}_{11}$				$\widehat{\Lambda}_{11}$			
(T_1, T_2) Marginal Survival Rates	Λ_{11}	Sample Mean	Sample SD	Empirical SD	95% CI Coverage	Sample Mean	Sample SD	Empirical SD	95% CI Coverage	Sample Mean	Sample SD	Empirical SD	95% CI Coverage
						z = 0							
(0.85, 0.85)	0.060	0.060	0.021	0.020	0.916	0.060	0.015	0.014	0.921	0.060	0.016	0.015	0.925
(0.85, 0.70)	0.114	0.115	0.041	0.039	0.910	0.115	0.029	0.028	0.934	0.113	0.025	0.024	0.925
(0.85, 0.55)	0.161	0.159	0.070	0.060	0.855	0.163	0.048	0.045	0.898	0.159	0.033	0.032	0.916
(0.70, 0.70)	0.226	0.226	0.083	0.076	0.900	0.226	0.055	0.055	0.927	0.224	0.041	0.040	0.932
(0.70, 0.55)	0.330	0.326	0.151	0.115	0.848	0.331	0.103	0.090	0.901	0.327	0.058	0.055	0.917
(0.55, 0.55)	0.500	0.483	0.274	0.162	0.750	0.497	0.198	0.139	0.843	0.499	0.082	0.078	0.935
						z = 1							
(0.85, 0.85)	0.046	0.044	0.016	0.015	0.903	0.045	0.011	0.011	0.927	0.046	0.017	0.017	0.928
(0.85, 0.70)	0.092	0.090	0.027	0.027	0.900	0.091	0.020	0.019	0.928	0.092	0.026	0.026	0.946
(0.85, 0.55)	0.140	0.138	0.043	0.042	0.876	0.139	0.031	0.030	0.921	0.140	0.034	0.035	0.947
(0.70, 0.70)	0.189	0.187	0.050	0.049	0.900	0.186	0.034	0.034	0.926	0.188	0.041	0.041	0.937
(0.70, 0.55)	0.290	0.290	0.082	0.076	0.906	0.287	0.055	0.054	0.927	0.288	0.054	0.054	0.943
(0.55, 0.55)	0.453	0.452	0.129	0.120	0.880	0.447	0.089	0.085	0.924	0.449	0.076	0.074	0.937

*
Sample mean and standard deviation (SD) based on 1000 simulations at each sample configuration. Empirical SD is the average of SD estimates based on the empirical variance estimator for $\widehat{\Lambda}_{11}$, and 95% CI coverage is the fraction of the 1000 simulated samples where the asymptotic confidence interval using the empirical SD includes the true Λ_{11} value.

Table 3:

Sample means and standard deviations for Volterra (V) and Dabrowska (D) estimators of the bivariate survival function under a Clayton–Oakes model (9) with $\gamma = 0$ with unit exponential marginal distributions and dependency parameter θ. Censoring times were also exponentially distributed and independent of each other with censoring rate c taking values 0 (no censoring), 2 (moderate censoring), or 5 (heavy censoring) at various sample sizes (n), based on 1000 simulations at each configuration

Marginal Survival Probabilities		Joint Survival Probability	Volterra Dabrowska	n = 100 c = 0	n = 100 c = 2	n = 100 c = 5	n = 250 c = 5	n = 500 c = 5
				$\theta = 0$ (independence)				
0.85	0.85	0.723	V	0.723(0.045)	0.723(0.050)	0.723(0.045)	0.722(0.036)	0.722(0.018)
			D	0.723(0.045)	0.723(0.050)	0.723(0.045)	0.722(0.036)	0.722(0.018)
0.85	0.70	0.595	V	0.594(0.048)	0.595(0.060)	0.594(0.048)	0.593(0.054)	0.595(0.026)
			D	0.594(0.048)	0.594(0.060)	0.594(0.048)	0.593(0.053)	0.595(0.026)
0.85	0.55	0.468	V	0.467(0.050)	0.468(0.072)	0.467(0.050)	0.465(0.089)	0.468(0.043)
			D	0.467(0.050)	0.467(0.071)	0.467(0.050)	0.465(0.095)	0.468(0.041)
0.70	0.70	0.490	V	0.491(0.050)	0.491(0.067)	0.491(0.050)	0.492(0.069)	0.489(0.050)
			D	0.491(0.050)	0.490(0.066)	0.491(0.050)	0.431(0.069)	0.488(0.047)
0.70	0.55	0.385	V	0.385(0.048)	0.386(0.078)	0.385(0.048)	0.394(0.102)	0.381(0.075)
			D	0.385(0.048)	0.385(0.073)	0.385(0.048)	0.387(0.123)	0.3880(0.086)
0.55	0.55	0.303	V	0.302(0.044)	0.306(0.090)	0.302(0.044)	0.311(0.122)	0.298(0.095)
			D	0.302(0.044)	0.304(0.085)	0.302(0.044)	0.316(0.142)	0.303(0.117)
				$\theta = 2$ (cross ratio= 3)				
0.85	0.85	0.752	V	0.752(0.043)	0.752(0.047)	0.753(0.054)	0.752(0.034)	0.752(0.025)
			D	0.752(0.043)	0.752(0.047)	0.753(0.054)	0.752(0.034)	0.752(0.025)
0.85	0.70	0.642	V	0.641(0.048)	0.641(0.058)	0.642(0.084)	0.641(0.053)	0.642(0.037)
			D	0.641(0.048)	0.641(0.058)	0.642(0.087)	0.641(0.052)	0.642(0.036)
0.85	0.55	0.521	V	0.519(0.050)	0.519(0.072)	0.529(0.135)	0.520(0.087)	0.520(0.064)
			D	0.519(0.050)	0.519(0.070)	0.522(0.156)	0.519(0.090)	0.520(0.063)
0.70	0.70	0.570	V	0.569(0.049)	0.570(0.064)	0.572(0.112)	0.565(0.070)	0.571(0.050)
			D	0.569(0.049)	0.570(0.062)	0.574(0.128)	0.567(0.070)	0.571(0.046)
0.70	0.55	0.480	V	0.478(0.049)	0.480(0.078)	0.493(0.150)	0.483(0.108)	0.479(0.086)
			D	0.478(0.049)	0.480(0.071)	0.480(0.183)	0.480(0.136)	0.480(0.084)

Marginal Survival Probabilities		Joint Survival Probability	Volterra Dabrowska	$n = 100$ $c = 0$	$n = 100$ $c = 2$	$n = 100$ $c = 5$	$n = 250$ $c = 5$	$n = 500$ $c = 5$
0.55	0.55	0.422	V	0.422(0.048)	0.426(0.095)	0.429(0.171)	0.416(0.138)	0.412(0.111)
			D	0.422(0.048)	0.427(0.086)	0.427(0.196)	0.421(0.160)	0.422(0.118)

Table 4:

Simulation summary statistics [a] for confidence intervals (CIs) for bivariate survival probability F at selected marginal survival probabilities under the Clayton–Oakes model (9) with $\beta_{10} = \beta_{01} = \gamma = \log 2$ and dependency parameter $\theta = 2$, at both $z = 0$ and $z = 1$.

Sample size (n)		500				1000				250			
T1 and T2 Failing %		22.6				22.6				100			
		$\hat{F}$				$\hat{F}$				$\hat{F}$			
(T_1, T_2) Marginal Survival Rates	F	Sample Mean	Sample SD	Bootstrap SD	95% CI Coverage	Sample Mean	Sample SD	Bootstrap SD	95% CI Coverage	Sample Mean	Sample SD	Empirical SD	95% CI Coverage
							$z = 0$						
(0.85, 0.85)	0.752	0.752	0.031	0.030	0.933	0.752	0.021	0.021	0.943	0.754	0.028	0.028	0.936
(0.85, 0.70)	0.642	0.644	0.046	0.044	0.931	0.644	0.030	0.031	0.950	0.644	0.035	0.034	0.940
(0.85, 0.55)	0.521	0.526	0.070	0.071	0.945	0.523	0.048	0.048	0.948	0.522	0.040	0.038	0.930
(0.70, 0.70)	0.570	0.570	0.057	0.059	0.960	0.571	0.040	0.040	0.961	0.571	0.037	0.036	0.933
(0.70, 0.55)	0.480	0.482	0.089	0.089	0.967	0.482	0.062	0.062	0.966	0.481	0.039	0.038	0.935
(0.55, 0.55)	0.422	0.417	0.142	0.123	0.951	0.423	0.0918	0.091	0.953	0.423	0.040	0.039	0.935
							$z = 1$						
(0.85, 0.85)	0.739	0.739	0.027	0.028	0.951	0.739	0.020	0.020	0.938	0.739	0.032	0.034	0.958
(0.85, 0.70)	0.623	0.623	0.036	0.036	0.943	0.623	0.025	0.025	0.951	0.624	0.037	0.038	0.960
(0.85, 0.55)	0.501	0.501	0.045	0.046	0.953	0.503	0.032	0.032	0.953	0.503	0.039	0.040	0.956
(0.70, 0.70)	0.538	0.538	0.041	0.042	0.953	0.538	0.029	0.029	0.949	0.540	0.039	0.040	0.952
(0.70, 0.55)	0.445	0.445	0.050	0.051	0.959	0.445	0.035	0.035	0.953	0.447	0.040	0.040	0.957
(0.55, 0.55)	0.379	0.378	0.063	0.064	0.965	0.379	0.043	0.044	0.957	0.380	0.040	0.040	0.947

[a] Sample mean and standard deviation (SD) based on 1000 simulations at each sample configuration. Bootstrap SD is the SD estimates from averaging the sample variances for $\hat{F}$ based on 200 bootstrap replicates for each generated sample, and 95% CI coverage is the fraction of the 1000 simulated samples where the asymptotic confidence interval using the bootstrap SD includes the true F value.

Table 5:

Simulation summary statistic[a] for bootstrap-based rectangular confidence bands for Λ_{11} and F at $z=0$ and $z=1$, using supremum statistics. Samples were generated under the Clayton–Oakes model (9) with $\beta_{10}=\beta_{01}=\gamma=\log 2$ and dependency parameter $\theta=2$, with z a binary variate taking values 0 and 1 with probability 0.5.

Sample size (n)		1000			250		
T_1 and T_2 Failure%		22.6			100		
		Λ_{11}			Λ_{11}		
(T_1,T_2) Marginal Survival Rates	Λ_{11}	Mean Boot. $C_{0.95}(\widehat{\Lambda}_{11})$	SD Boot. $C_{0.95}(\widehat{\Lambda}_{11})$	95% CB Coverage	Mean Boot. $C_{0.95}(\widehat{\Lambda}_{11})$	SD Boot. $C_{0.95}(\widehat{\Lambda}_{11})$	95% CB Coverage
$z=0$							
(0.85, 0.85)[b]	0.060	0.95	0.19	0.916	0.51	0.11	0.921
(0.70, 0.70)	0.226	3.61	0.92	0.917	1.36	0.23	0.934
(0.55, 0.55)	0.500	8.47	3.46	0.837	2.60	0.41	0.920
$z=1$							
(0.85, 0.85)	0.046	0.75	0.14	0.915	0.51	0.11	0.921
(0.70, 0.70)	0.189	2.34	0.39	0.920	1.36	0.23	0.934
(0.55, 0.55)	0.453	5.75	1.28	0.923	2.60	0.41	0.920
		F			F		
(T_1, T_2) Marginal Survival Rates	F	Mean Boot. $C_{0.95}(\widehat{F})$	SD Boot. $C_{0.95}(\widehat{F})$	95% CB Coverage	Mean Boot. $C_{0.95}(\widehat{F})$	SD Boot. $C_{0.95}(\widehat{F})$	95% CB Coverage
$z=0$							
(0.85, 0.85)	0.752	1.53	0.12	0.939	1.02	0.10	0.931
(0.70, 0.70)	0.570	2.89	0.26	0.961	1.38	0.10	0.937
(0.55, 0.55)	0.422	5.89	1.80	0.928	1.60	0.10	0.932
$z=1$							
(0.85, 0.85)	0.739	1.53	0.11	0.942	1.29	0.12	0.952
(0.70, 0.70)	0.538	2.32	0.16	0.952	1.66	0.12	0.946
(0.55, 0.55)	0.379	3.34	0.31	0.947	1.83	0.12	0.942

Author Manuscript

[a]Mean Boot $C_{0.95}(\widehat{\Lambda}_{11})$ is the average of the 95th percentiles of $W_{\widehat{\Lambda}_{11}}(t_1, t_2)$ over the 1000 samples, each with 200 bootstrap samples, and SD Boot $C_{0.95}(\widehat{\Lambda}_{11})$ is the standard deviation of these bootstrap percentiles; 95% CB Coverage is the fraction of bootstrap-based 95% confidence bands that include $\Lambda_{11}(s_1, s_2)$ for all $(s_1, s_2) \epsilon [0, t_1] \times [0, t_2]$. The same quantities are presented for confidence bands for F.

[b]Confidence region over $[0, -\log(0.85)] \times [0, -\log(0.85)]$.

Table 6:

Joint survival probability estimators ($\widehat{F}$) for breast cancer incidence and total mortality in the Women's Health Initiative Dietary Modification Trial ($n =$ 48,835, with 1764 women with incident breast cancers, 2508 deaths, and 134 women with both breast cancer and death during the 8.5 year average trial intervention period).

Follow-up Years for Breast Cancer (T_1)		Comparison Group ($z = 0$) Follow-up Years for Mortality (T_2)			Intervention Group ($z = 1$) Follow-up Years for Mortality (T_2)		
		3	6	9	3	6	9
3	$\widehat{F}$	0.979	0.961	0.936	0.980	0.962	0.937
	95% CI[a]	(0.978,0.980)	(0.959,0.963)	(0.933,0.939)	(0.979,0.981)	(0.960,0.965)	(0.934,0.941)
	95% CB[b]	(0.975,0.983)	(0.957,0.965)	(0.932,0.940)	(0.975,0.985)	(0.957,0.968)	(0.932,0.942)
6	$\widehat{F}$	0.965	0.947	0.923	0.967	0.949	0.925
	95% CI	(0.962,0.967)	(0.945,0.950)	(0.919,0.926)	(0.965,0.969)	(0.947,0.952)	(0.921,0.929)
	95% CB	(0.961,0.968)	(0.943,0.951)	(0.919,0.926)	(0.962,0.972)	(0.944,0.954)	(0.920,0.930)
9	$\widehat{F}$	0.951	0.933	0.909	0.954	0.937	0.912
	95% CI	(0.948,0.953)	(0.930,0.936)	(0.905,0.913)	(0.951,0.957)	(0.933,0.940)	(0.908,0.917)
	95% CB	(0.947,0.955)	(0.929,0.937)	(0.905,0.913)	(0.949,0.959)	(0.932,0.942)	(0.907,0.917)

[a]95% confidence intervals for F given z based on 200 bootstrap replicates.

[b]95% supremum-type confidence bands for F given z over the region $[0, 9] \times [0, 9]$ years, based on 200 bootstrap replicates.

Table 7:

Simulations summary statistics for survival function estimators $\hat{F}$ at $z = 0$ and $z = 1$ at selected marginal survival function percentiles under Clayton–Oakes model (20) with $\beta_{10} = 0$, $\beta_{01} = \log 2$, $\theta = 2$ and $\gamma = \log 2$.

Sample size (n)		1000	250	1000	250
T_1 and T_2 failure %		18.2	100	18.2	100
		$\hat{F}$–no time-varying covariates		$\hat{F}$–time-varying covariates included	
(T_1, T_2) Percentiles	F	Mean (SD)[a]	Mean (SD)	Mean (SD)	Mean (SD)
z = 0					
(0.85,0.85)	0.752	0.751(0.021)	0.747(0.030)	0.752(0.022)	0.754(0.036)
(0.85,0.70)	0.642	0.640(0.029)	0.632(0.035)	0.641(0.030)	0.650(0.039)
(0.85,0.55)	0.521	0.517(0.048)	0.506(0.039)	0.519(0.051)	0.532(0.042)
(0.70,0.70)	0.570	0.571(0.041)	0.553(0.037)	0.571(0.042)	0.569(0.041)
(0.70,0.55)	0.480	0.475(0.065)	0.453(0.039)	0.476(0.069)	0.483(0.042)
(0.55,0.55)	0.422	0.421(0.094)	0.392(0.039)	0.420(0.106)	0.417(0.041)
z = 1					
(0.85,0.85)	0.739	0.739(0.021)	0.743(0.031)	0.739(0.022)	0.741(0.034)
(0.85,0.70)	0.623	0.624(0.026)	0.632(0.035)	0.624(0.026)	0.624(0.039)
(0.85,0.55)	0.501	0.504(0.033)	0.514(0.038)	0.503(0.033)	0.499(0.041)
(0.70,0.70)	0.538	0.537(0.034)	0.550(0.037)	0.538(0.034)	0.546(0.039)
(0.70,0.55)	0.445	0.446(0.042)	0.462(0.038)	0.447(0.043)	0.449(0.040)
(0.55,0.55)	0.379	0.374(0.064)	0.396(0.038)	0.378(0.068)	0.389(0.040)

[a] Sample mean and standard deviation (SD) based on 1000 simulated samples at each sampling configuration.