Chapter 21
Semiparametric Analysis of Treatment Effect via Failure Probability Ratio and the Ratio of Cumulative Hazards

Song Yang

21.1 Introduction

For clinical trials with time to event data, often proportional hazards (Cox 1972) is assumed when comparing two treatment arms, and a single value of the hazard ratio is used to describe the group difference. When the proportionality assumption may not hold true, a natural approach to assess the time-dependency of the treatment effect is to analyze the hazard ratio function. For example, a conventional method is to give a hazard ratio estimate over each of a few time intervals, by fitting a piece-wise proportional hazards model. Alternatively, a "defined" time-varying covariate can be used in a Cox regression model, resulting in a parametric form for the hazard ratio function (e.g., Kalbfleisch and Prentice 2002, Chap. 6). With these approaches, it may not be easy to pre-specify the partition of the time axis or the parametric form of the hazard ratio function. Also, although the hazard ratio provides a nice display of temporal pattern of the treatment effect, it may not directly translate to the survival experience. It is possible for the hazard ratio to be less than 1 in a region where there is no improvement in the survival probability, or more than 1 in a region where the survival probability is not reduced. Similar phenomena also exists for the average of hazard ratio. Thus to assess the cumulative treatment effect, other measures can be used to supplement the hazard ratio.

Let $F_T(t)$ and $F_C(t)$ be the cumulative distribution functions of the two comparison groups, named treatment and control, respectively. The failure probability ratio

$$RR(t) = \frac{F_T(t)}{F_C(t)}$$

is the process version of relative risk, a measure often used in epidemiology. It directly indicates if the failure probability in the time interval $(0, t]$ is lower in the treatment group than in the control group, regardless of the possible up and down pattern

S. Yang (✉)

Office of Biostatistics Research, National Heart, Lung, and Blood Institute,
6701 Rockledge Dr. MSC 7913, Bethesda, MD, 20892, USA
e-mail: yangso@nhlbi.nih.gov

of the hazard ratio within $(0, t]$. Let $\Lambda_T(t)$ and $\Lambda_C(t)$ be the cumulative distribution functions of the two comparison groups respectively. The ratio of cumulative hazards

$$CHR(t) = \frac{\Lambda_T(t)}{\Lambda_C(t)}$$

also indicates the cumulative treatment effect, taking value < 1 if and only if $F_T(t) < F_C(t)$. Unlike the failure probability ratio, a value 0.8 for the ratio of cumulative hazards does not translate to a 20 % reduction of the failure probability. However, there is a nice property that if one adopts a proportional hazards adjustment for baseline covariates, then the ratio of cumulative hazards remains the same while the failure probability ratio depends on those covariates.

Although measures such as the failure probability ratio and the ratio of cumulative hazards provide usual supplementary information in addition to the hazard ratio, and the non-parametric estimators are easily available via the Nelson–Aalen estimator for the cumulative hazard function (Nelson 1969; Aalen 1975) and the Kaplan–Meier estimator of the survival function (Kaplan and Meier 1958), the non-parametric inference procedures are not used frequently, as the estimates are often not very smooth and the confidence intervals can be quite wide near the beginning of the data range. In this chapter, we consider semiparametric inference on the two ratios under a sufficiently flexible model. Assume that the failure times are absolutely continuous. The short-term and long-term hazards model proposed in Yang and Prentice (2005) postulates that

$$\lambda_T(t) = \frac{1}{e^{-\beta_2} + (e^{-\beta_1} - e^{-\beta_2})S_C(t)}\lambda_C(t), \ t < \tau_0, \tag{21.1}$$

where β_1, β_2 are scalar parameters, S_C the survivor function of the control group, $\lambda_T(t)$, $\lambda_C(t)$ the hazard function of the two groups respectively, and

$$\tau_0 = \sup\left\{x : \int_0^x \lambda_C(t)dt < \infty\right\}. \tag{21.2}$$

Under this model, $\lim_{t\downarrow 0} \lambda_T(t)/\lambda_C(t) = e^{\beta_1}$, $\lim_{t\uparrow\tau_0} \lambda_T(t)/\lambda_C(t) = e^{\beta_2}$. Thus, various patterns of the hazard ratio can be realized, including proportional hazards, no initial effect, disappearing effect, and crossing hazards. In particular, model (21.1) includes the proportional hazards model and the proportional odds model as special cases. There is no need to pre-specify a partition of the time axis or a parametric form of the hazard ratio function. For this model, Yang and Prentice (2005) proposed a pseudo-likelihood method for estimating the parameters, and Yang and Prentice (2011) studied inference procedures on the hazard ratio function and the average of the hazard ratio function. Extension of model (21.1) to the regression setting was also studied for current status data in Tong et al. (2007).

In the sections to follow, we first obtain the estimates and point-wise confidence intervals of the two ratios under model (21.1). Since the ratios are functions of time, simultaneous confidence intervals, or confidence bands, of the ratios are more

appropriate than the point-wise confidence intervals. We will employ a resampling scheme to obtain the confidence bands of the ratios. Such semiparametric inference procedures are applicable in a wide range of applications due to the properties of model (21.1) mentioned before. They will be illustrated through applications to data from two clinical trials.

Some previous work is related to the problems considered here. Dong and Matthews (2012) developed empirical likelihood estimator for the ratio of cumulative hazards with covariate adjustment. Schaubel and Wei (2011) considered several measures under dependent censoring and non-proportional hazards, and point-wise confidence intervals were constructed. In earlier works, Dabrowska et al. (1989) introduced a relative change function defined in terms of cumulative hazards and found simultaneous bands for this function under the assumption of proportional hazards. Parzen et al. (1997) constructed nonparametric simultaneous confidence bands for the survival probability difference. Cheng et al. (1997) proposed point-wise and simultaneous confidence interval procedures for the survival probability under semiparametric transformation models. McKeague and Zhao (2002) proposed simultaneous confidence bands for ratios of survival functions via the empirical likelihood method.

The article is organized as follows. In Sect. 21.2 the short-term and long-term hazard ratio model and the parameter estimator are described. Point-wise confidence intervals are established for the failure probability ratio and the ratio of cumulative hazards. In Sect. 21.3, confidence bands are developed. Simulation results are presented in Sect. 21.4. Applications to data from two clinical trials are given in Sect. 21.5. Some discussion is given in Sect. 21.6.

21.2   The Estimators and Point-Wise Confidence Intervals

Denote the pooled lifetimes of the two groups by $T_1, \cdots, T_n$, with $T_1, \cdots, T_{n_1}$, $n_1 < n$, constituting the control group. Let $C_1, \cdots, C_n$ be the censoring variables, and $Z_i = I(i > n_1)$, $i = 1, \cdots, n$, where $I(\cdot)$ is the indicator function. The available data consist of the independent triplets (X_i, δ_i, Z_i), $i = 1, \ldots, n$, where $X_i = \min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$. We assume that T_i, C_i are independent given Z_i. The censoring variables (C_i's) need not be identically distributed, and in particular the two groups may have different censoring patterns. For $t < \tau_0$ with τ_0 defined in (21.2), let $R(t)$ be the odds function $1/S_C(t) - 1$ of the control group. The model of Yang and Prentice (2005) can be expressed as

$$\lambda_i(t) = \frac{1}{e^{-\beta_1 Z_i} + e^{-\beta_2 Z_i} R(t)} \frac{dR(t)}{dt}, \ i = 1, \ldots, n, t < \tau_0,$$

where $\lambda_i(t)$ is the hazard function for T_i given Z_i.

Under model (21.1), $RR(t)$ and $CHR(t)$ depends on the parameter $\beta = (\beta_1, \beta_2)^T$ and the baseline function $R(t)$, where "T" denotes transpose. Yang and Prentice (2005) studied a pseudo likelihood estimator $\hat{\beta}$ of β which we describe below.

Let $\tau < \tau_0$ be such that

$$\lim_n \sum_{i=1}^n I(X_i \geq \tau) > 0, \tag{21.3}$$

with probability 1. For $t \leq \tau$, define

$$\hat{P}(t; \mathbf{b}) = \prod_{s \leq t} \left(1 - \frac{\sum_{i=1}^n \delta_i e^{-b_2 Z_i} I(X_i = s)}{\sum_{i=1}^n I(X_i \geq s)} \right),$$

$$\hat{R}(t; \mathbf{b}) = \frac{1}{\hat{P}(t; \mathbf{b})} \int_0^t \frac{\hat{P}_-(s; \mathbf{b})}{\sum_{i=1}^n I(X_i \geq s)} d \left(\sum_{i=1}^n \delta_i e^{-b_1 Z_i} I(X_i \leq s) \right),$$

where $\hat{P}_-(s; \mathbf{b})$ denotes the left continuous (in s) version of $\hat{P}(s; \mathbf{b})$. Let $L(\beta, R)$ be the likelihood function of β under model (21.1) when the function $R(t)$ is known, with the corresponding score vector $S(\beta, R) = \partial \ln L(\beta, R)/\partial \beta$. Define $Q(\mathbf{b}) = S(\mathbf{b}, R)|_{R(t) = \hat{R}(t; \mathbf{b})}$. Then the pseudo maximum likelihood estimator $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)^T$ of β is the zero of $Q(\mathbf{b})$. Note that the use of $\hat{R}(t; \mathbf{b})$ results in the estimating function $Q(\mathbf{b})$ which does not involve the infinite dimensional nuisance parameter $R(t)$, thus the finite dimensional parameter β can be estimated much more easily.

Once $\hat{\beta}$ is obtained, $R(t)$ can be estimated by $\hat{R}(t; \hat{\beta})$. Thus under model (21.1), plugging-in the estimators $\hat{\beta}$ and $\hat{R}(t; \hat{\beta})$, we can estimate the failure probability ratio $RR(t)$ and the ratio of cumulative hazards $CHR(t)$ by

$$\widehat{RR}(t) = \frac{1 + \hat{R}(t; \hat{\beta})}{\hat{R}(t; \hat{\beta})} \left(1 - \left\{ 1 + e^{-\hat{\beta}_2 + \hat{\beta}_1} \hat{R}(t; \hat{\beta}) \right\}^{-e^{\hat{\beta}_2}} \right), \tag{21.4}$$

and

$$\widehat{CHR}(t) = \frac{e^{\hat{\beta}_2} \ln \left\{ 1 + e^{-\hat{\beta}_2 + \hat{\beta}_1} \hat{R}(t; \hat{\beta}) \right\}}{\ln \left\{ 1 + \hat{R}(t; \hat{\beta}) \right\}}, \tag{21.5}$$

respectively. Note that under the model and with the pseudo likelihood estimator, the distributions of the two groups share a common baseline function $R(t)$ which is estimated using pooled data. Thus the resulting estimators for $RR(t)$ and $CHR(t)$ are expected to be smoother and more stable than the nonparametric estimators. In Appendix A, we show that, under certain regularity conditions, the two estimators in (21.4) and (21.5) are strongly consistent under model (21.1). To study the distributional properties of the estimators, let

$$U_n(t) = \sqrt{n}(\widehat{RR}(t) - RR(t)), \ t \leq \tau,$$

$$V_n(t) = \sqrt{n}(\widehat{CHR}(t) - CHR(t)), \ t \leq \tau,$$

and

$$\Omega = \left\{ -\frac{1}{n}\frac{\partial Q(\beta)}{\partial \beta}\right\}^{-1}.$$

Let $\hat{\Omega}$ be an estimator of Ω defined by replacing β with $\hat{\beta}$ and $R(t)$ with $\hat{R}(t;\hat{\beta})$.

In Appendix B we show that, for $t \leq \tau$, the processes U_n and V_n are asymptotically equivalent to, respectively,

$$
\begin{aligned}
\tilde{U}_n(t) = {} & \frac{A_{RR}^T(t)\Omega}{\sqrt{n}}\left(\sum_{i \leq n_1}\int_0^\tau \mu_1 dM_i + \sum_{i > n_1}\int_0^\tau \mu_2 dM_i\right) \\
& + \frac{B_{RR}(t)}{\sqrt{n}}\left(\sum_{i \leq n_1}\int_0^t v_1 dM_i + \sum_{i > n_1}\int_0^t v_2 dM_i\right)
\end{aligned}
\tag{21.6}
$$

and

$$
\begin{aligned}
\tilde{V}_n(t) = {} & \frac{A_{CHR}^T(t)\Omega}{\sqrt{n}}\left(\sum_{i \leq n_1}\int_0^\tau \mu_1 dM_i + \sum_{i > n_1}\int_0^\tau \mu_2 dM_i\right) \\
& + \frac{B_{CHR}(t)}{\sqrt{n}}\left(\sum_{i \leq n_1}\int_0^t v_1 dM_i + \sum_{i > n_1}\int_0^t v_2 dM_i\right),
\end{aligned}
\tag{21.7}
$$

where A_{RR}, A_{CHR}, μ_1, μ_2 are appropriately defined 2×1 vector functions and B_{RR}, B_{CHR}, v_1, v_2 scalar functions given in Appendix B. It will then be shown that U_n and V_n converge weakly to some zero-mean Gaussian processes U^* and V^* respectively. With estimators $\hat{B}_{RR}(t)$, $\hat{A}_{RR}(t)$, ... , given in Appendix B, it will be shown that the limiting covariance functions of U^* and V^* can be consistently estimated, respectively, by

$$
\begin{aligned}
\hat{\sigma}_{RR}(s,t) = {} & \hat{A}_{RR}^T(s)\hat{\Omega}\left(\int_0^\tau \frac{\hat{\mu}_1(w)\hat{\mu}_1^T(w)K_1(w)d\hat{R}(w;\hat{\beta})}{n(1+\hat{R}(w;\hat{\beta}))}\right. \\
& \left. + \int_0^\tau \frac{\hat{\mu}_2(w)\hat{\mu}_2^T(w)K_2(w)d\hat{R}(w;\hat{\beta})}{n(e^{-\hat{\beta}_1}+e^{-\hat{\beta}_2}\hat{R}(w;\hat{\beta}))}\right)\hat{\Omega}^T\hat{A}_{RR}(t) \\
& + \hat{B}_{RR}(s)\hat{B}_{RR}(t)\left(\int_0^s \frac{\hat{v}_1^2(w)K_1(w)d\hat{R}(w;\hat{\beta})}{n(1+\hat{R}(w;\hat{\beta}))}\right. \\
& \left. + \int_0^s \frac{\hat{v}_2^2(w)K_2(w)d\hat{R}(w;\hat{\beta})}{n(e^{-\hat{\beta}_1}+e^{-\hat{\beta}_2}\hat{R}(w;\hat{\beta}))}\right) \\
& + \hat{B}_{RR}(t)\hat{A}_{RR}^T(s)\hat{\Omega}\left(\int_0^t \frac{\hat{\mu}_1(w)\hat{v}_1(w)K_1(w)d\hat{R}(w,\hat{\beta})}{n(1+\hat{R}(w;\hat{\beta}))}\right.
\end{aligned}
$$

$$+ \int_0^t \frac{\hat{\mu}_2(w)\hat{v}_2(w)K_2(w)d\hat{R}(w,\hat{\beta})}{n(e^{-\hat{\beta}_1} + e^{-\hat{\beta}_2}\hat{R}(w;\hat{\beta}))}\Bigg)$$

$$+ \hat{B}_{RR}(s)\hat{A}_{RR}^T(t)\hat{\Omega}\left(\int_0^s \frac{\hat{\mu}_1(w)\hat{v}_1(w)K_1(w)d\hat{R}(w,\hat{\beta})}{n(1 + \hat{R}(w;\hat{\beta}))}\right.$$

$$\left. + \int_0^s \frac{\hat{\mu}_2(w)\hat{v}_2(w)K_2(w)d\hat{R}(w,\hat{\beta})}{n(e^{-\hat{\beta}_1} + e^{-\hat{\beta}_2}\hat{R}(w;\hat{\beta}))}\right), \tag{21.8}$$

and

$$\hat{\sigma}_{CHR}(s,t) = \hat{A}_{CHR}^T(s)\hat{\Omega}\left(\int_0^\tau \frac{\hat{\mu}_1(w)\hat{\mu}_1^T(w)K_1(w)d\hat{R}(w;\hat{\beta})}{n(1 + \hat{R}(w;\hat{\beta}))}\right.$$

$$\left. + \int_0^\tau \frac{\hat{\mu}_2(w)\hat{\mu}_2^T(w)K_2(w)d\hat{R}(w;\hat{\beta})}{n(e^{-\hat{\beta}_1} + e^{-\hat{\beta}_2}\hat{R}(w;\hat{\beta}))}\right)\hat{\Omega}^T\hat{A}_{CHR}(t)$$

$$+ \hat{B}_{CHR}(s)\hat{B}_{CHR}(t)\left(\int_0^s \frac{\hat{v}_1^2(w)K_1(w)d\hat{R}(w;\hat{\beta})}{n(1 + \hat{R}(w;\hat{\beta}))}\right.$$

$$\left. + \int_0^s \frac{\hat{v}_2^2(w)K_2(w)d\hat{R}(w;\hat{\beta})}{n(e^{-\hat{\beta}_1} + e^{-\hat{\beta}_2}\hat{R}(w;\hat{\beta}))}\right)$$

$$+ \hat{B}_{CHR}(t)\hat{A}_{CHR}^T(s)\hat{\Omega}\left(\int_0^t \frac{\hat{\mu}_1(w)\hat{v}_1(w)K_1(w)d\hat{R}(w,\hat{\beta})}{n(1 + \hat{R}(w;\hat{\beta}))}\right.$$

$$\left. + \int_0^t \frac{\hat{\mu}_2(w)\hat{v}_2(w)K_2(w)d\hat{R}(w,\hat{\beta})}{n(e^{-\hat{\beta}_1} + e^{-\hat{\beta}_2}\hat{R}(w;\hat{\beta}))}\right)$$

$$+ \hat{B}_{CHR}(s)\hat{A}_{CHR}^T(t)\hat{\Omega}\left(\int_0^s \frac{\hat{\mu}_1(w)\hat{v}_1(w)K_1(w)d\hat{R}(w,\hat{\beta})}{n(1 + \hat{R}(w;\hat{\beta}))}\right.$$

$$\left. + \int_0^s \frac{\hat{\mu}_2(w)\hat{v}_2(w)K_2(w)d\hat{R}(w,\hat{\beta})}{n(e^{-\hat{\beta}_1} + e^{-\hat{\beta}_2}\hat{R}(w;\hat{\beta}))}\right). \tag{21.9}$$

The estimators $\hat{\Omega}$, $\hat{A}_{RR}(t)$, $\hat{A}_{CHR}(t)$ involve the derivative vector $\partial\hat{R}(t;\beta)/\partial\beta$ and the derivative matrix in Ω. From various simulation studies, these derivatives can be approximated by numerical derivatives for easier calculation, and the results are fairly stable with respect to the choice of the jump size in the numerical derivatives.

For a fixed $t_0 \leq \tau$, confidence intervals for $RR(t_0)$ can be obtained from the asymptotic normality of $\widehat{RR}(t_0)$ and the estimated variance $\hat{\sigma}_{RR}(t_0, t_0)$. For better small sample behavior and to ensure that the confidence intervals remain on the positive side of the axis as usual, we make a logarithm transformation resulting in the asymptotic $100(1 - \alpha)\%$ confidence interval

$$\widehat{RR}(t_0) \exp\left(\mp z_{\alpha/2} \frac{\sqrt{\hat{\sigma}_{RR}(t_0, t_0)}}{\sqrt{n}\widehat{RR}(t_0)}\right), \tag{21.10}$$

where $z_{\alpha/2}$ is the $100(1-\alpha/2)\%$ percentile of the standard normal distribution.

Similarly, for $CHR(t_0)$, an asymptotic $100(1-\alpha)\%$ confidence interval is

$$\widehat{CHR}(t_0) \exp\left(\mp z_{\alpha/2} \frac{\sqrt{\hat{\sigma}_{CHR}(t_0, t_0)}}{\sqrt{n}\widehat{CHR}(t_0)}\right). \tag{21.11}$$

21.3 Confidence Bands

For simultaneous inference on $RR(t)$ over a time interval $I = [a, b] \subset [0, \tau]$, let $w_n(t)$ be a data-dependent function that converges in probability to a bounded function $w^*(t) > 0$, uniformly in t over I. Then, it follows that U_n/w_n converges weakly U^*/w^*. Let c_α be the upper αth percentile of $\sup_{t \in I} |U^*/w^*|$, then an asymptotic $100(1-\alpha)\%$ simultaneous confidence band for $RR(t)$, $t \in I$, can be obtained as

$$\widehat{RR}(t) \exp\left(\mp c_\alpha \frac{w_n(t)}{\sqrt{n}\widehat{RR}(t)}\right). \tag{21.12}$$

The analytic form of c_α is quite intractable. The bootstrapping method provides a well established alternative approach. However, it is very time-consuming. More discussion on this is described further on the applications to clinical trial data in Sect. 21.5. Here we have used a normal resampling approximation similar to the approach used in Lin et al. (1993). This approach results in substantial savings in computing time, and has been used in many works, including Lin et al. (1994), Cheng et al. (1997), Tian et al. (2005), and Peng and Huang (2007).

For $t \le \tau$, let $N_i(t) = \delta_i I(X_i \le t)$, $i = 1, \cdots, n$, and define the process

$$
\begin{aligned}
\hat{U}_n(t) &= \frac{\hat{A}_{RR}^T(t)\hat{\Omega}}{\sqrt{n}} \left(\sum_{i \le n_1} \int_0^\tau \hat{\mu}_1 d(\epsilon_i N_i) + \sum_{i > n_1} \int_0^\tau \hat{\mu}_2 d(\epsilon_i N_i) \right) \\
&\quad + \frac{\hat{B}_{RR}(t)}{\sqrt{n}} \left(\sum_{i \le n_1} \int_0^t \hat{v}_1 d(\epsilon_i N_i) + \sum_{i > n_1} \int_0^t \hat{v}_2 d(\epsilon_i N_i) \right) \\
&= \frac{\hat{A}_{RR}^T(t)\hat{\Omega}}{\sqrt{n}} \left(\sum_{i \le n_1} \epsilon_i \delta_i \hat{\mu}_1(X_i) I(X_i \le \tau) + \sum_{i > n_1} \epsilon_i \delta_i \hat{\mu}_2(X_i) I(X_i \le \tau) \right) \\
&\quad + \frac{\hat{B}_{RR}(t)}{\sqrt{n}} \left(\sum_{i \le n_1} \epsilon_i \delta_i \hat{v}_1(X_i) I(X_i \le t) + \sum_{i > n_1} \epsilon_i \delta_i \hat{v}_2(X_i) I(X_i \le t) \right),
\end{aligned}
\tag{21.13}
$$

where ϵ_i, $i = 1, \ldots, n$, are independent standard normal variables that are also independent of the data. Conditional on (X_i, δ_i, Z_i), $i = 1, \ldots, n$, $\hat{U}_n$ is a sum of n independent variables at each time point. In Appendix B, it will be shown that $\hat{U}_n$ given the data converges weakly to U^*. It follows that $\sup_{t \in I} |\hat{U}_n(t)/w_n(t)|$ given the data converges in distribution to $\sup_{t \in I} |U^*(t)/w^*(t)|$. Therefore, c_α can be estimated empirically from a large number of realizations of the conditional distribution of $\sup_{t \in I} |\hat{U}/w|$ given the data.

Similarly, to considerations in Yang and Prentice (2011) for inference on the hazard ratio, we look at several choices of the weight w_n. For $w_n(t) = \sqrt{\hat{\sigma}_{RR}(t, t)}$ we obtain the equal precision bands (Nair 1984), which only differ from point-wise confidence intervals in using c_α instead of $z_{\alpha/2}$. For $w_n(t) = 1 + \hat{\sigma}_{RR}(t, t)$ we obtain the Hall–Wellner type bands recommended by Bie et al. (1987). The simplest case $w_n(t) \equiv 1$ does not require the computation of $\hat{\sigma}_{RR}(t, t)$, and hence is easier to implement.

To obtain simultaneous confidence bands for $CHR(t)$, let

$$
\begin{aligned}
\hat{V}_n(t) &= \frac{\hat{A}_{CHR}^T(t)\hat{\Omega}}{\sqrt{n}} \left(\sum_{i \leq n_1} \int_0^\tau \hat{\mu}_1 d(\epsilon_i N_i) + \sum_{i > n_1} \int_0^\tau \hat{\mu}_2 d(\epsilon_i N_i) \right) \\
&\quad + \frac{\hat{B}_{CHR}(t)}{\sqrt{n}} \left(\sum_{i \leq n_1} \int_0^t \hat{v}_1 d(\epsilon_i N_i) + \sum_{i > n_1} \int_0^t \hat{v}_2 d(\epsilon_i N_i) \right) \\
&= \frac{\hat{A}_{CHR}^T(t)\hat{\Omega}}{\sqrt{n}} \left(\sum_{i \leq n_1} \epsilon_i \delta_i \hat{\mu}_1(X_i) I(X_i \leq \tau) + \sum_{i > n_1} \epsilon_i \delta_i \hat{\mu}_2(X_i) I(X_i \leq \tau) \right) \\
&\quad + \frac{\hat{B}_{CHR}(t)}{\sqrt{n}} \left(\sum_{i \leq n_1} \epsilon_i \delta_i \hat{v}_1(X_i) I(X_i \leq t) + \sum_{i > n_1} \epsilon_i \delta_i \hat{v}_2(X_i) I(X_i \leq t) \right),
\end{aligned}
$$
(21.14)

where ϵ_i, $i = 1, \ldots, n$, are independent standard normal variables that are also independent of the data. Let $\tilde{w}_n(t)$ be a data-dependent function that converges in probability to a bounded function $\tilde{w}^*(t) > 0$, uniformly in t over I. Let $\tilde{c}_\alpha$ be upper αth percentile of $\sup_{t \in [a,b]} |V^*(t)/\tilde{w}^*|$. Similarly, to the argument above for $RR(t)$, an asymptotic $100(1 - \alpha)\%$ simultaneous confidence band for $CHR(t)$, $t \in I$, can be obtained as

$$
\widehat{CHR}(t) \exp \left(\mp \tilde{c}_\alpha \frac{\tilde{w}_n(t)}{\sqrt{n}\widehat{CHR}(t)} \right),
$$
(21.15)

where $\tilde{c}_\alpha$ can be approximated empirically from a large number of realizations of the conditional distribution of $\sup_{t \in [a,b]} |\hat{V}(t)/\tilde{w}_n|$ given the data. For $\tilde{w}_n = \sqrt{\hat{\sigma}_{CHR}(t, t)}, 1 + \hat{\sigma}_{CHR}(t, t)$ and $\tilde{w}_n \equiv 1$ respectively, we obtain the equal precision, Hall–Wellner type, and unweighted confidence bands for $CHR(t)$.

21.4 Simulation Studies

For stable moderate sample behavior, we restrict the range of the confidence bands for both $RR(t)$ and $CHR(t)$. The range is between the 40th percentile of the uncensored data at the lower end and the 95th percentile of the uncensored data at the upper end. The lower end point of this range seems a little high compared to other situations such as the inference on the hazard ratio in Yang and Prentice (2011). This is to provide a range in which the nonparametric procedures and the proposed model-based procedures in (21.10–21.12) and (21.15) are to be compared. Toward the beginning of the data range, the nonparametric estimates can be very unstable and the confidence intervals can be quite wide, as will be illustrated in the data example to follow. Also, compared with the hazard ratio as a measure of the temporal pattern of the treatment effect, $RR(t)$ and $CHR(t)$ measure the cumulative treatment effect. Thus in biomedical research, there is little interest in their behaviors near the beginning of the data range. In various applications to clinical trial data, the specified range for the confidence bands is not nearly as restrictive as it seems and contains a meaningful interval of the data range. In the estimating procedures, the function $\hat{P}(t; \mathbf{b})$ is replaced by an asymptotically equivalent form

$$\exp\left(-\int_0^t \frac{1}{\sum_{i=1}^n I(X_i \geq s)} d\left\{\sum_{i=1}^n \delta_i e^{-b_2 Z_i} I\left(X_i \leq s\right)\right\}\right).$$

For simulation studies reported here and for the real data application in Sect. 21.5, τ was set to include all data in calculating $\hat{\beta}$. All numerical computations were done in *Matlab*. Some representative results of simulation studies are given in Table 21.1, where lifetime variables were generated with $R(t)$ chosen to yield the standard exponential distribution for the control group. The values of β were $(\log(.9), \log(1.2))$ and $(\log(1.2), \log(.8))$, representing 1/3 increase or decrease over time from the initial hazard ratio, respectively. The censoring variables were independent and identically distributed with the log-normal distribution, where the normal distribution had mean c and standard deviation 0.5, with c chosen to achieve various censoring rates. The data were split into the treatment and control groups by a 1:1 ratio. The empirical coverage probabilities were obtained from 1000 repetitions, and for each repetition, the critical values c_α and $\tilde{c}_\alpha$ were calculated empirically from 1,000 realizations of relevant conditional distributions. For both $RR(t)$ and $CHR(t)$, the equal precision bands, Hall–Wellner type bands and unweighted bands are denoted by EP, HW, and UW respectively.

Note that with 1,000 repetitions and $1.96\sqrt{.95 \cdot 0.05/1000} = 0.0135$, we expect the empirical coverage probabilities to be mostly greater than 0.9365. In Table 21.1, for *RR*, the empirical coverage probabilities are greater than 0.9365 for all but one case with the smallest sample size $n = 100$ and at 50 % censoring. For *CHR*, the confidence bands are mostly conservative, with all empirical coverage probabilities greater than 0.95. One plausible explanation for this conservative phenomenon could be that the estimate for $CHR(t)$ is more directly related to the martingales associated with censored data, resulting in better approximations.

Table 21.1 Empirical coverage probabilities of the three types of confidence bands HW, EP, and UW, for the failure probability ratio RR and the ratio of cumulative hazards CHR, under model (21.1), based on 1000 repetitions

Hazard ratio	Censoring	n	RR			CHR		
			HW	EP	UW	HW	EP	UW
0.9 ↑ 1.2	10 %	100	0.949	0.967	0.958	0.977	0.983	0.986
	30 %		0.967	0.966	0.962	0.976	0.984	0.989
	50 %		0.943	0.957	0.966	0.969	0.970	0.975
	10 %	200	0.950	0.973	0.965	0.972	0.978	0.985
	30 %		0.964	0.969	0.968	0.959	0.970	0.986
	50 %		0.959	0.968	0.974	0.969	0.975	0.980
	10 %	400	0.960	0.974	0.971	0.958	0.969	0.982
	30 %		0.961	0.970	0.971	0.959	0.968	0.981
	50 %		0.958	0.974	0.977	0.967	0.975	0.983
1.2 ↓ 0.8	10 %	100	0.946	0.963	0.962	0.974	0.980	0.980
	30 %		0.958	0.964	0.968	0.969	0.977	0.981
	50 %		0.926	0.940	0.966	0.962	0.969	0.974
	10 %	200	0.958	0.967	0.952	0.968	0.975	0.971
	30 %		0.958	0.959	0.958	0.972	0.974	0.968
	50 %		0.946	0.954	0.961	0.962	0.964	0.974
	10 %	400	0.960	0.957	0.954	0.969	0.972	0.973
	30 %		0.960	0.969	0.960	0.966	0.969	0.970
	50 %		0.949	0.962	0.959	0.967	0.974	0.972

21.5   Applications

For the Women's Health Initiative (WHI) randomized controlled trial of combined (estrogen plus progestin) postmenopausal hormone therapy, an elevated coronary heart disease risk was reported, with overall unfavorable health benefits versus risks over an average of 5.6-year study period (Writing Group 2002; Manson et al. 2003). After controlling for time from estrogen-plus-progestin initiation and confounding, hazard ratio estimates still indicate elevated risk of coronary heart disease and venous thromboembolism early on during the trial, under a piece-wise Cox model assuming constant hazard ratio separately on 0–2 years, 2–5 years, and 5+ years (Prentice et al. 2005). Let us first illustrate the methods developed in the previous sections with the venous thromboembolism (VTE) data from the WHI clinical trial. Among the 16,608 postmenopausal women ($n_1 = 8102$), there were 167 and 76 events observed in the treatment and control group respectively, implying about 98.5 % censoring, primarily by the trial stopping time. Fitting model (21.1) to this data set, we get $\hat{\beta} = (4.72, 0.014)^T$. Plots of the model based survival curves and the Kaplan–Meier curves for the two groups show that the model is reasonable. For $RR(t)$, the three 95 % simultaneous confidence bands (EP, HW, and UW) under model (21.1) are given in Fig. 21.1, together with the point estimates. The nonparametric point estimates are also included to compare with the model-based estimates. Furthermore, model-based 95 % point-wise confidence intervals are included as well, to indicate by how much the confidence intervals are widened to improve from point-wise to
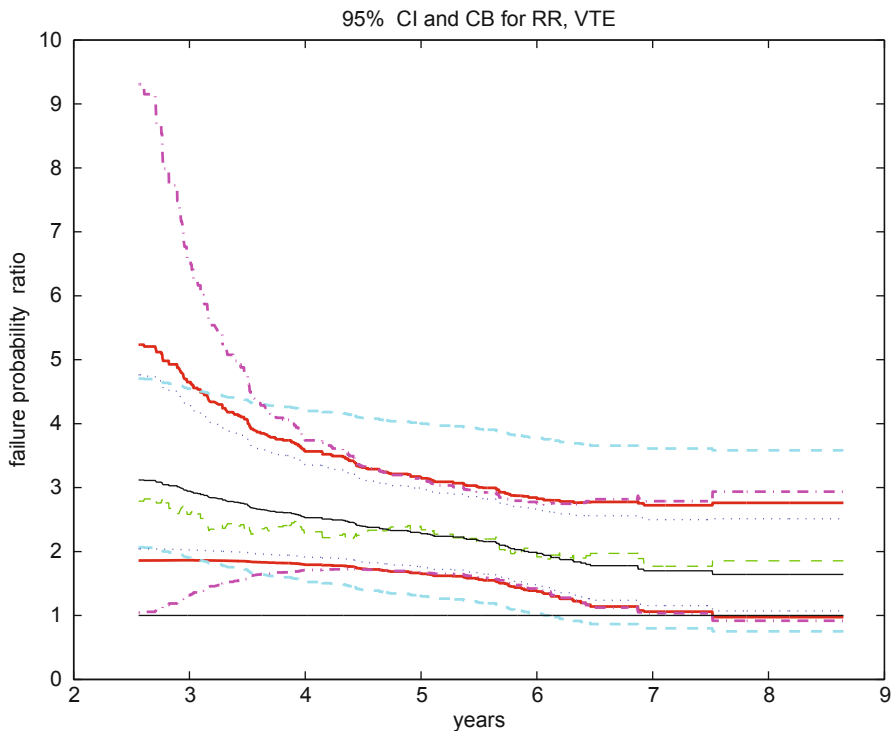
Fig. 21.1 95 % point-wise confidence intervals and simultaneous confidence bands of the failure probability ratio for the WHI VTE data: *Outside red solid lines*—equal precision confidence band, *magenta dash-dotted lines*—Hall–Wellner confidence band, *outside cyan dashed lines*—unweighted confidence band, *dotted lines*—95 % point-wise confidence intervals, *central black solid line*—the estimated failure probability ratio under the model, *central green dashed line*—the estimated failure probability ratio using Kaplan–Meier estimators

simultaneous coverage. From Fig. 21.1, it can be seen that the Hall–Wellner type band and the equal precision band are almost the same a little after the 4th year. However, the Hall–Wellner type band is noticeably wider toward the beginning of the date range. The unweighted band maintains a roughly constant width through the data range considered, which is roughly as wide as the equal precision band at the begining of the data range, but wider throughout the rest of the data range. Similar phenomena are often seen in additional applications not reported here. Based on various applications and simulation studies, we recommend that the equal precision band be used in making inference on $RR(t)$ under model (21.1).

For $CHR(t)$, the 95 % point-wise confidence intervals and confidence bands under model (21.1) are given in Fig. 21.2. Similarly to the case for $RR(t)$, the equal precision band is preferred in making inference on $CHR(t)$ under model (21.1). From Fig. 21.1 and 21.2, there is evidence that from 2.5 to 7.5 years, the event probability is higher in the treatment group than in the control group.
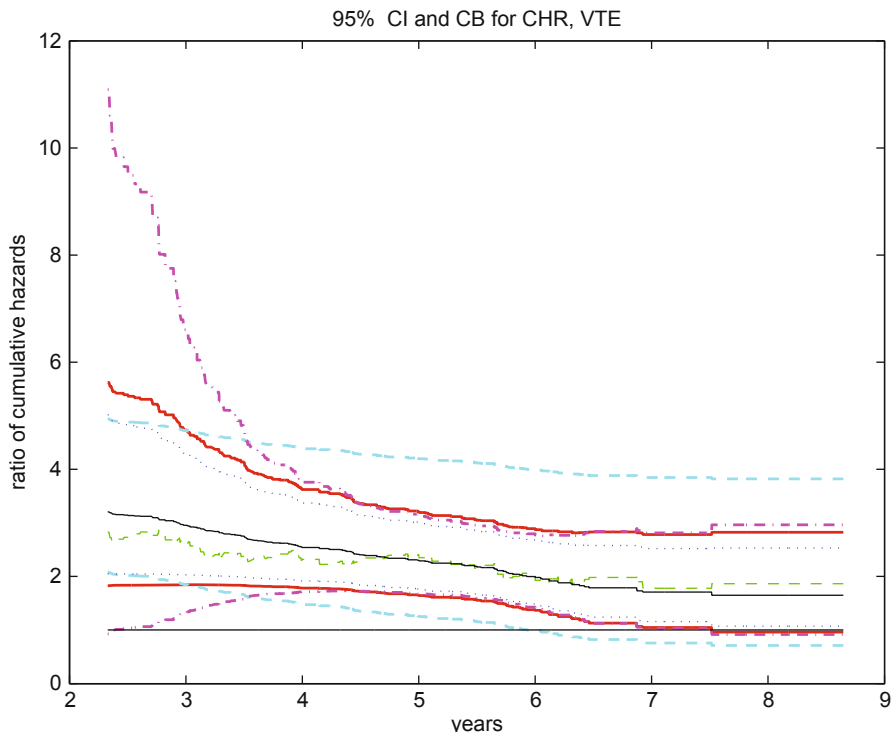
Fig. 21.2 95 % point-wise confidence intervals and simultaneous confidence bands of the ratio of cumulative hazards for the WHI VTE data: *Outside red solid lines*—equal precision confidence band, *magenta dash-dotted lines*—Hall–Wellner confidence band, em outside cyan dashed lines—unweighted confidence band, *dotted lines*—95 % point-wise confidence intervals, *central black solid line*—the estimated failure probability ratio under the model, *central green dashed line*—the estimated failure probability ratio using Kaplan–Meier estimators

For comparison, from Yang and Prentice (2011), the 95 % point-wise confidence intervals and equal precision confidence band are obtained for the hazard ratio under model (21.1), given in Fig. 21.3. The results are in good agreement with the results under the piece-wise Cox model used in Prentice et al. (2005). In an interval near the beginning of the data range, there is greater hazard of venous thromboembolism in the treatment group than in the control group. This interval has shorter length than the intervals in Fig. 21.1 and 21.2 where the treatment group has a higher event probability than in the control group.

Note that the simple bootstrap method for approximating c_α and $\tilde{c}_\alpha$, when $w_n \equiv 1$ and $\tilde{w}_n \equiv 1$ respectively, is already much more computationally intensive than the normal resampling approximation employed here. With $w_n(t) = \sqrt{\hat{\sigma}_{RR}(t)}$ and $\tilde{w}_n = \sqrt{\hat{\sigma}_{CHR}(t)}$, the bootstrap method would require one more level of bootstrapping samples to obtain the estimated variance functions, thus further increasing the computational burden. In comparison, once $\hat{\sigma}_{RR}(t)$ and $\hat{\sigma}_{CHR}(t)$ are obtained, the
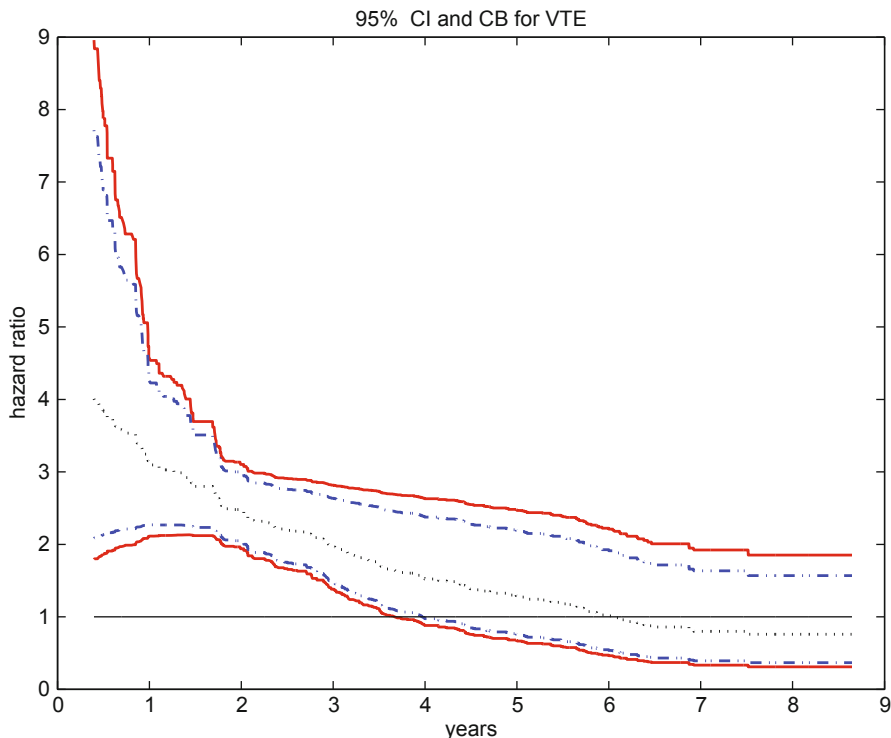
Fig. 21.3 95 % point-wise confidence intervals and simultaneous confidence bands of the hazard ratio function for the WHI VTE data: *Red solid lines*—equal precision confidence band, *blue dash-dotted lines*—95 % point-wise confidence intervals, *dotted line*—the estimated hazard ratio function

normal resampling approximation only needs a small additional computation and programming cost.

To see how the nonparametric procedures compare with the proposed model-based procedures, Fig. 21.4 presents 95 % point-wise confidence intervals, both model-based and nonparametric, together with the point estimates, of $CHR(t)$ for the VTE data from WHI. It can be seen that the nonparametric estimates and confidence intervals can be quite unstable near the beginning of the data range. As $t \downarrow 0$, the hazard ratio at t and $CHR(t)$ should both approach the same limit, which is e^{β_1} under the model. From Fig. 21.4, the model-based estimator of $CHR(t)$ near $t = 0$ takes values around 5, which is comparable to results in the literature, while the nonparametric estimator of $CHR(t)$ near $t = 0$ takes much more extreme values. Also, the model-based estimates and confidence intervals are smoother throughout, and the confidence intervals are often narrower than their nonparametric counterparts. Similar phenomena are also present for $RR(t)$ (omitted). This is a major reason that the nonparametric estimates for $RR(t)$ and $CHR(t)$ are rarely used in biomedical studies.
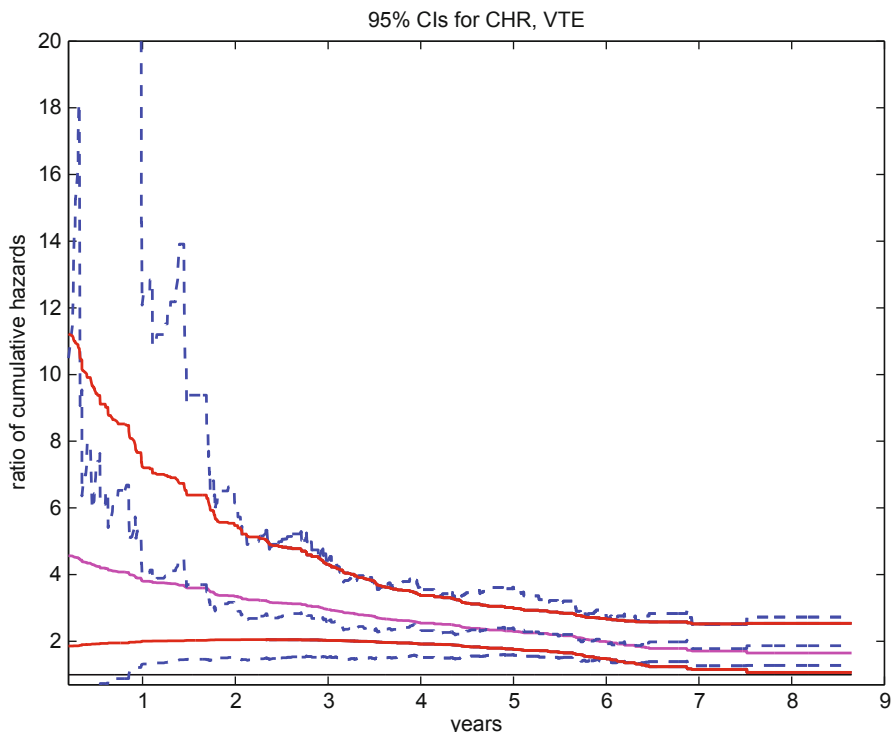
Fig. 21.4 Model-based and nonparametric 95 % point-wise confidence intervals of the ratio of cumulative hazards for the WHI VTE data: *Outside red solid lines*—model based 95 % point-wise confidence intervals, *outside blue dashed lines*—nonparametric 95 % point-wise confidence intervals, *central magenta solid line*—model based estimate of the ratio of cumulative hazards, *central blue dashed line*—nonparametric estimate of the ratio of cumulative hazards

Next, we look at an example with mild violation of the proportional hazards assumption. The Digoxin Intervention Group trial (The Digitalis investigation group 1997) was a randomized, double-blind clinical trial on the effect of digoxin on mortality and hospitalization. In the main trial, patients with left ventricular ejection fraction of 0.45 or less were randomized to digoxin (3397 patients) or placebo (3403 patients) in addition to diuretics and angiotensin-converting-enzyme inhibitors. We look at the data on death attributed to worsening heart failure. For testing the proportional hazards assumption, the acceleration test statistic of Breslow et al. (1984) gives a *p*- value of 0.098. This indicates some mild proportionality violation. For $RR(t)$, the 95 % point-wise confidence intervals and confidence bands under model (21.1) are given in Fig. 21.4. Possibly due to only a mild violation of the proportionality assumption, the Hall–Wellner type band, the equal precision band and the unweighted band are almost the same for the entire data range considered. From Fig. 21.4, there is evidence that for the range of 1.5–3 year, the treatment reduces the event probability.
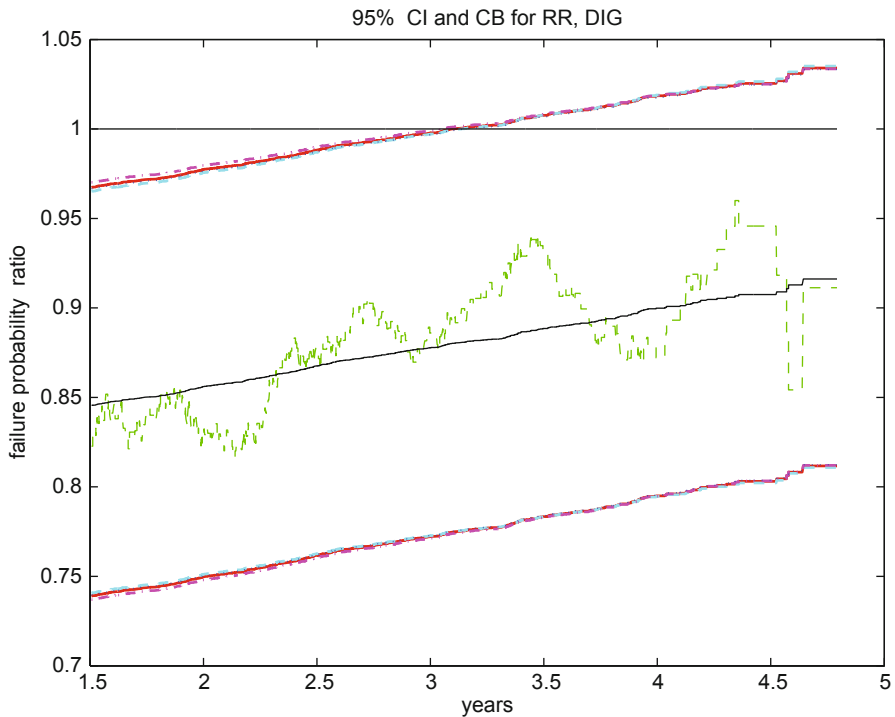
Fig. 21.5 95 % point-wise confidence intervals and simultaneous confidence bands of the failure probability ratio for the DIG data: *Outside red solid lines*—equal precision confidence band, *magenta dash-dotted lines*—Hall–Wellner confidence band, *outside cyan dashed lines*—unweighted confidence band, *Dotted lines*—95 % point-wise confidence intervals, *central black solid line*—the estimated failure probability ratio under the model, *central green dashed line*—the estimated failure probability ratio using Kaplan–Meier estimators

For *CHR(t)*, the 95 % point-wise confidence intervals and confidence bands under model (21.1) are given in Fig. 21.5. Again all three confidence bands are very close to each other. From Fig. 21.5, there is evidence of reduced event probability in the treatment group for the range of 1.3 year to 3 years.

Again for comparison, from Yang and Prentice (2011), the 95 % point-wise confidence intervals and equal precision confidence band are obtained for the hazard ratio under model (21.1), given in Fig. 21.6. From Fig. 21.6, there is evidence that from 0 to .75 year, in the treatment group there is reduced hazard of death attributed to worsening heart failure. Note that this range is much narrower than the range where there is evidence of reduced event probability in the treatment group seen from Fig. 21.4 and 21.5 (Fig. 21.7).
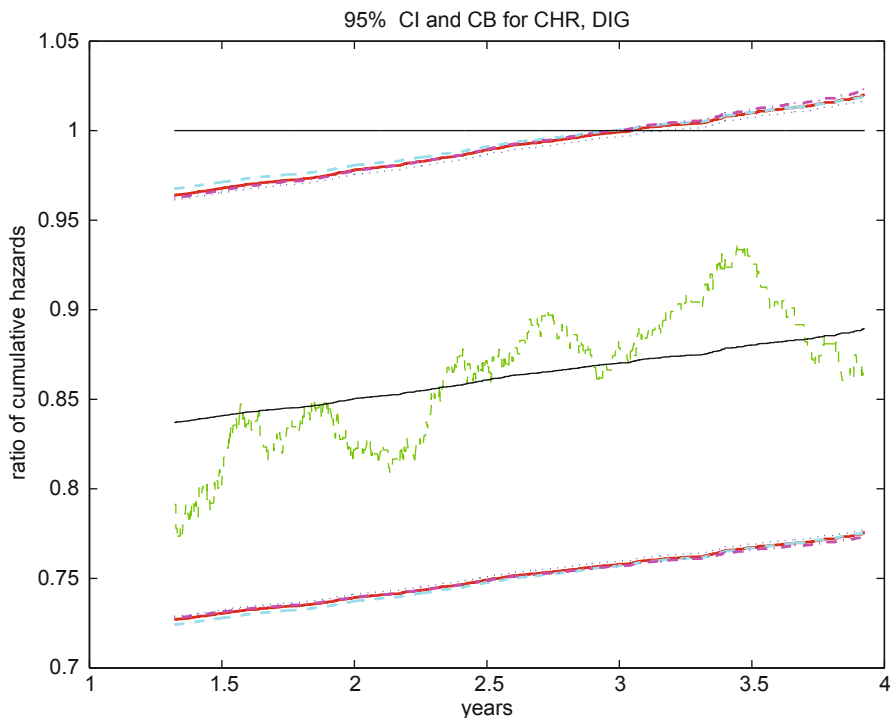
Fig. 21.6 95 % point-wise confidence intervals and simultaneous confidence bands of the ratio of cumulative hazards for the DIG data: *Outside red solid lines*—equal precision confidence band, *magenta dash-dotted lines*—Hall–Wellner confidence band, *outside cyan dashed lines*—unweighted confidence band, *dotted lines*—95 % point-wise confidence intervals, *central black solid line*—the estimated failure probability ratio under the model, *central green dashed line*—the estimated failure probability ratio using Kaplan–Meier estimators

21.6 Discussion

We have studied the asymptotic properties of the estimators for the failure probability ratio and the ratio of cumulative hazards under a semiparametric model applicable to a sufficiently wide range of applications. Point-wise confidence intervals and confidence bands are developed for the two ratios. In simulation studies, the confidence bands have good performance for moderate samples. Among the confidence bands with different weights, the equal precision confidence band is recommended based on various simulation studies and clinical trial data applications. Similarly, inference procedures can be developed for the odds ratio. The point-wise confidence intervals and confidence bands for the odds ratio are usually wider than the corresponding intervals and bands for the failure probability ratio and the ratio of cumulative hazards. Due to space limit those results are not presented here. When the censoring is heavy, there are very little differences among the confidence intervals and bands for the failure probability ratio, the ratio of cumulative hazards, and the odds ratio. The
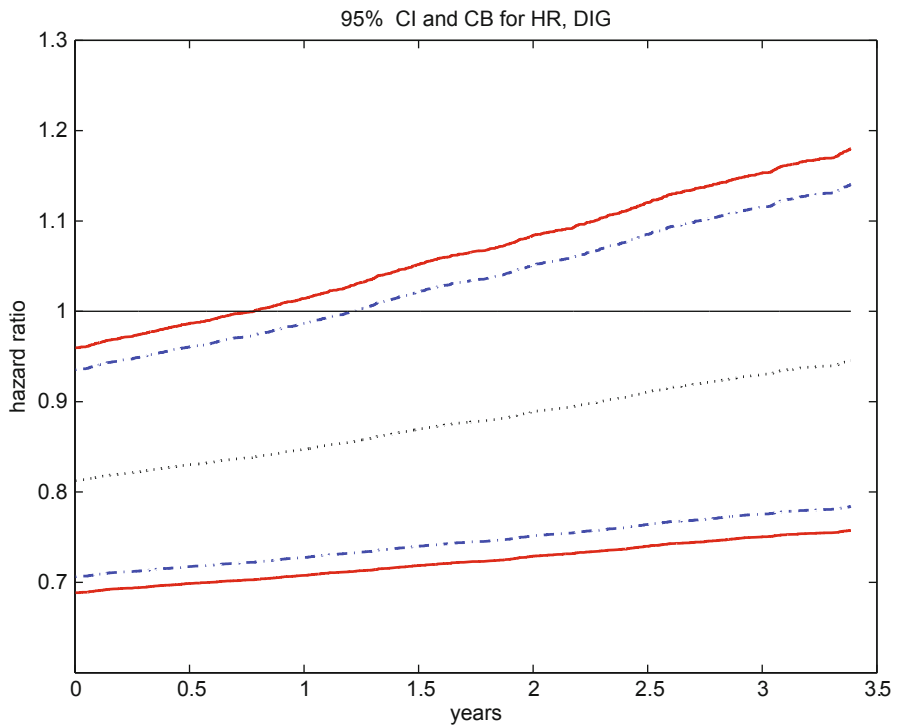
Fig. 21.7 95 % point-wise confidence intervals and simultaneous confidence bands of the hazard ratio function for the DIG data: *Red solid lines*—equal precision confidence band, *blue dash-dotted lines*—95 % point-wise confidence intervals, *dotted line*—the estimated hazard ratio function

confidence intervals and bands presented here provide good visual tools for assessing cumulative effect of the treatment. They can supplement the visual tools based on the hazard ratio which focuses the temporal pattern of the treatment effect. It is also of interest to extend the results here by considering adjustment for covariate via a regression analysis. These and other problems are worthy of further exploration.

Appendix A: Consistency

The following regularity conditions will be assumed throughout the Appendices:

Condition 1. $\lim \frac{n_1}{n} = \rho \in (0, 1)$.

Condition 2. The survivor function G_i of C_i given Z_i is continuous and satisfies

$$\frac{1}{n} \sum_{i \le n_1} G_i(t) \to \Gamma_1, \quad \frac{1}{n} \sum_{i > n_1} G_i(t) \to \Gamma_2,$$

uniformly for $t \leq \tau$, for some Γ_1, Γ_2, and $\tau < \tau_0$ such that $\Gamma_j(\tau) > 0$, $j = 1, 2$.
Condition 3. The survivor functions S_C and S_T are absolutely continuous and $S_C(\tau) > 0$.

Under these conditions, the strong law of large numbers implies that (21.3) is satisfied.

For $t \leq \tau$, define

$$L(t) = \Gamma_1 S_C + \Gamma_2 S_T,$$

$$U_j(t; \mathbf{b}) = \int_0^t \Gamma_1 dF_C + \exp\left(-b_j\right) \int_0^t \Gamma_2 dF_T, \ j = 1, 2,$$

$$\Lambda_j(t; \mathbf{b}) = \int_0^t \frac{dU_j(s; \mathbf{b})}{L(s)}, \ j = 1, 2,$$

$$P(t; \mathbf{b}) = \exp\{-\Lambda_2(t; \mathbf{b})\}, \ R(t; \mathbf{b}) = \frac{1}{P(t; \mathbf{b})} \int_0^t P(s; \mathbf{b}) d\Lambda_1(s; \mathbf{b}),$$

$$f_j^0(t; \mathbf{b}) = \frac{\exp\left(-b_j\right) R^{j-1}(t; \mathbf{b})}{\exp\left(-b_1\right) + \exp\left(-b_2\right) R(t; \mathbf{b})}, \ j = 1, 2,$$

$$m_j(\mathbf{b}) = \left\{ \int_0^\tau f_j^0 \Gamma_2(t) dF_T(t) - \int_0^\tau \frac{f_j^0 \Gamma_2(t) S_T(t) dR(t; \mathbf{b})}{\exp\left(-b_1\right) + \exp\left(-b_2\right) R(t; \mathbf{b})} \right\}, \ j = 1, 2,$$

and $m(\mathbf{b}) = \left(m_1(\mathbf{b}), m_2(\mathbf{b})\right)'$. We will also assume

Condition 4. The function $m(\mathbf{b})$ is non-zero for $b \in \mathcal{B} - \{\beta\}$, where $\mathcal{B}$ is a compact neighborhood of β.

Theorem 1. *Suppose that Conditions $1 \sim 4$ hold. Then, (i) the zero $\hat{\beta}$ of $Q(\mathbf{b})$ in $\mathcal{B}$ is strongly consistent for β; (ii) $\widehat{RR}(t)$ is strongly consistent for $RR(t)$, uniformly for $t \in [0, \tau]$, and $\widehat{CHR}(t)$ is strongly consistent for $CHR(t)$, uniformly on $t \in [0, \tau]$; (iii) $\hat{\Omega}$ converges almost surely to a limiting matrix Ω^*.*

Proof. Under Conditions $1 \sim 3$, the limit of $\sum_{i=1}^n I(X_i \geq t)/n$ is bounded away from zero on $t \in [0, \tau]$. Thus, it can be shown that, with probability 1,

$$\frac{\sum_{i=1}^n \delta_i e^{-b_j Z_i} I(X_i = t)}{\sum_{i=1}^n \delta_i I(X_i \geq t)} \rightarrow 0, \ j = 1, 2, \ |\Delta \hat{P}(t; \mathbf{b})| \rightarrow 0, \ |\Delta \hat{R}(t; \mathbf{b})| \rightarrow 0,$$

$$(21.16)$$

uniformly for $t \in [0, \tau]$ and $b \in \mathcal{B}$, where Δ indicates the jump of the function in t.

Define the martingale residuals

$$\hat{M}_i(t; \mathbf{b}) = \delta_i I(X_i \leq t) - \int_0^t I(X_i \geq s) \frac{\hat{R}(ds; \mathbf{b})}{e^{-b_1 Z_i} + e^{-b_2 Z_i} \hat{R}(s; \mathbf{b})}, \ 1 \leq i \leq n.$$

From (21.16) and the fundamental theorem of calculus, it can be shown that, with probability 1,

$$Q(\mathbf{b}) = \sum_{i=1}^{n} \int_0^{\tau} \left\{ f_i(t; \mathbf{b}) + o(1) \right\} \hat{M}_i(dt; \mathbf{b}), \qquad (21.17)$$

uniformly in $t \leq \tau$, $b \in \mathcal{B}$ and $i \leq n$, where $f_i = (f_{1i}, f_{2i})^T$, with

$$f_{1i}(t; \mathbf{b}) = \frac{Z_i e^{-b_1 Z_i}}{e^{-b_1 Z_i} + e^{-b_2 Z_i} \hat{R}(t; \mathbf{b})}, \quad f_{2i}(t; \mathbf{b}) = \frac{Z_i e^{-b_2 Z_i} \hat{R}(t; \mathbf{b})}{e^{-b_1 Z_i} + e^{-b_2 Z_i} \hat{R}(t; \mathbf{b})}.$$

From the strong law of large numbers (Pollard 1990, p. 41) and repeated use of Lemma A1 of Yang and Prentice (2005), one obtain, with probability 1,

$$\hat{P}(t; \mathbf{b}) \to \hat{P}(t; \mathbf{b}), \ \hat{R}(t; \mathbf{b}) \to R(t; \mathbf{b}), \ Q(\mathbf{b})/n \to m(\mathbf{b}), \qquad (21.18)$$

uniformly in $t \leq \tau$ and $\mathbf{b} \in \mathcal{B}$. From these results and Condition 4, one obtains the strong consistency of $\widehat{RR}(t)$ and $\widehat{CHR}(t)$, and almost sure convergence of $\hat{\Omega}$.

Appendix B: Weak Convergence

Let $\xi_0(t) = 1 + R(t), \xi(t) = e^{-\beta_1} + e^{-\beta_2} R(t), \hat{\xi}_0(t) = 1 + \hat{R}(t; \beta), \hat{\xi}(t) = e^{-\beta_1} + e^{-\beta_2} \hat{R}(t; \beta)$, and define

$$K_1(t) = \sum_{i \leq n_1} I(X_i \geq t), \ K_2(t) = \sum_{i > n_1} I(X_i \geq t),$$

$$H(t) = \frac{1}{\hat{\xi}(t)} (e^{-\beta_1}, e^{-\beta_2} \hat{R}(t; \beta))^T,$$

$$J(t) = \int_t^{\tau} \frac{H(s) K_1(s) K_2(s)}{\hat{\xi}(s) \hat{P}(s; \beta)} \left(\frac{e^{-\beta_2}}{\xi(s)} - \frac{1}{\xi_0(s)} \right) dR(s).$$

Similarly, to the proof of Theorem 1, it can be shown that, with probability 1,

$$Q(\beta) = \sum_{i \leq n_1} \int_0^{\tau} \{\mu_1(t) + o(1)\} dM_i(t) + \sum_{i > n_1} \int_0^{\tau} \{\mu_2(t) + o(1)\} dM_i(t), \quad (21.19)$$

uniformly in $t \leq \tau$ and $i \leq n$, where

$$\mu_1(t) = -\frac{\hat{\xi}_0(t) H(t) K_2(t)}{\hat{\xi}(t) K(t)} + \frac{\hat{\xi}_0(t) \hat{P}_-(t; \beta)}{K} J(t),$$

$$\mu_2(t) = H(t) \frac{K_1(t)}{K(t)} + \frac{\hat{\xi}(t) \hat{P}_-(t; \beta)}{K(t)} J(t), \qquad (21.20)$$

$$M_i(t) = \delta_i I(X_i \le t) - \int_0^t I(X_i \ge s) \frac{dR(s)}{e^{-\beta_1 Z_i} + e^{-\beta_2 Z_i} R(s)}, \quad i = 1, \ldots, n.$$

By Lemma A3 of Yang and Prentice (2005),

$$\sqrt{n}(\hat{R}(t; \beta) - R(t)) = \frac{1}{\sqrt{n}\hat{P}(t; \beta)} \left(\sum_{i \le n_1} \int_0^t v_1 dM_i + \sum_{i > n_1} \int_0^t v_2 dM_i \right) \quad (21.21)$$

where

$$v_1(t) = \frac{n\xi_0(t)\hat{P}_-(t; \beta)}{K(t)}, \quad v_2(t) = \frac{n\xi(t)\hat{P}_-(t; \beta)}{K(t)}.$$

Define

$$A_{RR}(t) = \left(\frac{\hat{S}_T(t)}{\hat{F}_C(t)\hat{\xi}(t)} - \frac{\hat{F}_T(t)\hat{S}_C^2(t)}{\hat{F}_C^2(t)} \right) \frac{\partial \hat{R}(t; \beta)}{\partial \beta} + \frac{\hat{S}_T(t)}{\hat{F}_C(t)} \left(\frac{R(t)}{\xi(t)}, \Lambda_T(t) - \frac{R(t)}{\xi(t)} \right)^T,$$

$$B_{RR}(t) = \frac{1}{\hat{P}(t; \beta)} \left(\frac{\hat{S}_T(t)}{\hat{F}_C(t)\hat{\xi}(t)} - \frac{\hat{F}_T(t)\hat{S}_C^2(t)}{\hat{F}_C^2(t)} \right),$$

$$A_{CHR}(t) = \left(\frac{1}{\Lambda_C(t)\hat{\xi}(t)} - \frac{\Lambda_T(t)\hat{S}_C(t)}{\Lambda_C^2(t)} \right) \frac{\partial \hat{R}(t; \beta)}{\partial \beta} + \frac{1}{\Lambda_C(t)} \left(\frac{R(t)}{\xi(t)}, \Lambda_T(t) - \frac{R(t)}{\xi(t)} \right)^T,$$

$$B_{CHR}(t) = \frac{1}{\hat{P}(t; \beta)} \left(\frac{1}{\Lambda_C(t)\hat{\xi}(t)} - \frac{\Lambda_T(t)\hat{S}_C(t)}{\Lambda_C^2(t)} \right).$$

For $A_{RR}(t)$, $B_{RR}(t)$, $A_{CHR}(t)$, $B_{CHR}(t)$, $mu_1(t), \mu_2(t), v_1(t)$, $v_2(t)$, let $A_{RR}^*(t)$, $B_{RR}^*(t), \ldots$ etc. be their almost sure limit. In addition, let L_j be the almost sure limit of K_j/n, $j = 1, 2$. For $0 \le s, t < \tau$, let

$$\sigma_{RR}(s, t)$$

$$= A_{RR}^{*T}(s)\Omega^* \left(\int_0^\tau \frac{\mu_1^*\mu_1^{*T}}{1 + R} L_1 dR + \int_0^\tau \frac{\mu_2^*\mu_2^{*T}}{e^{-\beta_1} + e^{-\beta_2} R} L_2 dR \right) \Omega^{*T} D^*(t)$$

$$+ B_{RR}^*(s)B_{RR}^*(t) \left(\int_0^s \frac{v_1^{*2}}{1 + R} L_1 dR + \int_0^s \frac{v_2^{*2}}{e^{-\beta_1} + e^{-\beta_2} R} L_2 dR \right)$$

$$+ B_{RR}^*(t)A_{RR}^{*T}(s)\Omega^* \left(\int_0^t \frac{\mu_1^* v_1^*}{1 + R} L_1 dR + \int_0^t \frac{\mu_2^* v_2^*}{e^{-\beta_1} + e^{-\beta_2} R} L_2 dR \right)$$

$$+ B_{RR}^*(s)A_{RR}^{*T}(t)\Omega^* \left(\int_0^s \frac{\mu_1^* v_1^*}{1 + R} L_1 dR + \int_0^s \frac{\mu_2^* v_2^*}{e^{-\beta_1} + e^{-\beta_2} R} L_2 dR \right), \quad (21.22)$$

and

$$\sigma_{CHR}(s, t)$$

$$= A_{CHR}^{*T}(s)\Omega^* \left(\int_0^\tau \frac{\mu_1^* \mu_1^{*T}}{1+R} L_1 dR + \int_0^\tau \frac{\mu_2^* \mu_2^{*T}}{e^{-\beta_1} + e^{-\beta_2} R} L_2 dR \right) \Omega^{*T} D^*(t)$$

$$+ B_{CHR}^*(s) B_{CHR}^*(t) \left(\int_0^s \frac{v_1^{*2}}{1+R} L_1 dR + \int_0^s \frac{v_2^{*2}}{e^{-\beta_1} + e^{-\beta_2} R} L_2 dR \right)$$

$$+ B_{CHR}^*(t) A_{CHR}^{*T}(s)\Omega^* \left(\int_0^t \frac{\mu_1^* v_1^*}{1+R} L_1 dR + \int_0^t \frac{\mu_2^* v_2^*}{e^{-\beta_1} + e^{-\beta_2} R} L_2 dR \right)$$

$$+ B_{CHR}^*(s) A_{CHR}^{*T}(t)\Omega^* \left(\int_0^s \frac{\mu_1^* v_1^*}{1+R} L_1 dR + \int_0^s \frac{\mu_2^* v_2^*}{e^{-\beta_1} + e^{-\beta_2} R} L_2 dR \right). \quad (21.23)$$

For $A_{RR}(t)$, $B_{RR}(t), \ldots$, etc. define corresponding estimator $\hat{B}_{RR}(t)$, $\hat{A}_{RR}(t)$, ... by replacing β with $\hat{\beta}$, $R(t)$ with $\hat{R}(t; \hat{\beta})$. Defince $\hat{\sigma}_{RR}(s,t)$ and $\hat{\sigma}_{CHR}(s,t)$ by repalcing $B_{RR}(t)$, $A_{RR}(t)$, $\mu_1(t)$, $\mu_2(t)$, $v_1(t)$, $v_2(t), \ldots$ in $\sigma_{RR}(s,t)$ and $\sigma_{CHR}(s,t)$ by $\hat{B}_{RR}(t)$, $\hat{A}_{RR}(t), \ldots$ etc.

Theorem 2. *Suppose that Conditions $1 \sim 4$ hold and that the matrix Ω^* is nonsingular. Then, (i) U_n is asymptotically equivalent to the process $\tilde{U}_n$ in (21.6) which converges weakly to a zero-mean Gaussian process U^* on $[0, \tau]$, with covariance function $\sigma_{RR}(s,t)$ in (21.22). $\sigma_{RR}(s,t)$ can be consistently estimated by $\hat{\sigma}_{RR}(s,t)$. In addition, $\hat{U}_n(s)$ given the data converges weakly to the same limiting process U^*. (ii) $V_n(t)$ is asymptotically equivalent to the process $\tilde{V}_n$ in (21.7) which converges weakly to a zero-mean Gaussian process V^* on $[0, \tau]$, with covariance function $\sigma_{CHR}(s,t)$ in (21.23). $\sigma_{CHR}(s,t)$ can be consistently estimated by $\hat{\sigma}_{CHR}(s,t)$. In addition, $\hat{V}_n(s)$ given the data converges weakly to the same limiting process V^*.*

Proof. (i) As in the proof for Theorem A2 (ii) in Yang and Prentice (2005), by the strong embedding theorem and (21.19), $Q(\beta)/\sqrt{n}$ can be shown to be asymptotically normal. Now Taylor series expansion of $Q(\mathbf{b})$ around β and the non-singularity of Ω^* imply that $\sqrt{n}(\hat{\beta} - \beta)$ is asymptotically normal. From the $\sqrt{n}$- boundedness of $\hat{\beta}$,

$$\sqrt{n}(\hat{R}(t; \hat{\beta}) - \hat{R}(t; \beta)) = \frac{\partial R(t; \beta)}{\partial \beta} \sqrt{n}(\hat{\beta} - \beta) + o_p(1),$$

uniformly in $t \leq \tau$. These results, some algebra and Taylor series expansion together show that U_n is asymptotically equivalent to $\tilde{U}_n$. Similarly, to the proof of the asymptotic normality of $Q(\beta)/\sqrt{n}$, one can show that $\tilde{U}_n$ converges weakly to a zero-mean Gaussian process. Denote the limiting process by U^*. From the martingale integral representation of $\tilde{U}_n$, it can be shown that the covariation process of U^* is given by $\sigma(s,t)$ in (21.22). The consistency of $\hat{\sigma}_{RR}(s,t)$ can be shown similarly to the proof of Theorem 1.

By checking the tightness condition and the convergence of the finite-dimensional distributions, it can be shown that $\hat{U}_n(s)$ given the data also converges weakly to U^*.

(ii) The assertions on V_n, $\tilde{V}_n$, etc. can be proved similarly to the case for U_n, $\tilde{U}_n$, etc. in (i).

References

Aalen OO (1975) Statistical inference for a family of counting processes. PhD thesis. University of California, Berkeley

Bie O, Borgan O, Liestøl lK (1987) Confidence intervals and confidence bands for the cumulative hazard rate function and their small-sample properties. Scand J Stat 14:221–233

Breslow N, Elder L, Berger L (1984). A two sample censored-data rank test for acceleration. *Biometrics* **40**, 1042–1069

Cheng SC, Wei LJ, Ying Z (1997) Predicting survival probabilities with semiparametric transformation models. J Am Stat Assoc 92:227–235

Cox DR (1972) Regression models and life-tables (with Discussion). J R Statist Soc B 34:187–220

Dabrowska DM, Doksum KA, Song J (1989) Graphical comparison of cumulative hazards for two populations. Biometrika 76:763–773

The Digitalis Investigation Group (1997) The effect of Digoxin on mortality and morbidity in patients with heart failure. N Engl J Med 336:525–533

Dong B, Matthews DE (2012) Empirical likelihood for cumulative hazard ratio estimation with covariate adjustment. Biometrics 68:408–418

Kalbfleisch JD, Prentice RL (2002) The statistical analysis of failure time data, 2nd ed. Wiley, New York

Kaplan E, Meier P (1958) Nonparametric estimation from incomplete observations. J Am Stat Assoc 53:457–481

Lin DY, Wei LJ, Ying Z (1993) Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika 80:557–572

Lin DY, Fleming TR, Wei LJ (1994) Confidence bands for survival curves under the proportional hazards model. Biometrika 81:73–81

Manson JE, Hsia J, Johnson KC, Rossouw JE, Assaf AR, Lasser NL, Trevisan M, Black HR, Heckbert SR, Detrano R, Strickland OL, Wong ND, Crouse JR, Stein E, Cushman M, for the Women'S Health Initiative Investigators (2003) Estrogen plus progestin and the risk of coronary heart disease. New Eng J Med 349:523–534

McKeague IW, Zhao Y (2002) Simultaneous confidence bands for ratios of survival functions via empirical likelihood. Stat Probabil Lett 60:405–415

Nair VN (1984) Confidence bands for survival functions with censored data: a comparative study. Technometrics 26:265–275

Nelson W (1969) Hazard plotting for incomplete failure data. J Qual Technol 1:27–52

Parzen MI, Wei LJ, Ying Z (1997) Simultaneous confidence intervals for the difference of two survival functions. Scand J Stat 24:309–314

Peng L, Huang Y (2007) Survival analysis with temporal covariate effects. Biometrika 94:719–733

Pollard D (1990) Empirical processes: theory and applications. Institute of Mathematical Statistics, Hayward

Prentice RL, Langer R, Stefanick ML, Howard BV, Pettinger M, Anderson G, Barad D, Curb JD, Kotchen J, Kuller L, Limacher M, Wactawski-Wende J, for the Women'S Health Initiative Investigators (2005) Combined postmenopausal hormone therapy and cardiovascular disease: toward resolving the discrepancy between observational studies and the women's health initiative clinical trial. Amer J of Epi 162:404–414

Schaubel DE, Wei G (2011) Double inverse-weighted estimation of cumulative treatment effects under nonproportional hazards and dependent censoring. Biometrics 67:29–38

Tian L, Zucker D, Wei LJ (2005) On the Cox model with time-varying regression coefficients. J Am Statist Assoc 100:172–183

Tong X, Zhu C, Sun J (2007) Semiparametric regression analysis of two-sample current status data, with applications to tumorigenicity experiments. Canad J Statist 35:575–584

Writing Group for the Women's Health Initiative Investigators (2002) Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the women's health initiative randomized controlled trial. J Amer Med Assoc 288:321–333

Yang S, Prentice RL (2005) Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data. Biometrika 92:1–17

Yang S, Prentice RL (2011) Estimation of the 2-sample hazard ratio function using a semiparametric model. Biostatistics 12:354–368