Lifetime Data Anal (2013) 19:100–116
DOI 10.1007/s10985-012-9234-4

Estimation of odds of concordance based on the Aalen additive model

Torben Martinussen · Christian Bressen Pipper

Received: 19 December 2011 / Accepted: 6 October 2012 / Published online: 27 October 2012

Abstract The Cox regression model is often used when analyzing survival data as it provides a convenient way of summarizing covariate effects in terms of relative risks. The proportional hazards assumption may not hold, however. A typical violation of the assumption is time-changing covariate effects. Under such scenarios one may use more flexible models but the results from such models may be complicated to communicate and it is desirable to have simple measures of a treatment effect, say. In this paper we focus on the odds-of-concordance measure that was recently studied by Schemper et al. (Stat Med 28:2473–2489, 2009). They suggested to estimate this measure using weighted Cox regression (WCR). Although WCR may work in many scenarios no formal proof can be established. We suggest an alternative estimator of the odds-of-concordance measure based on the Aalen additive hazards model. In contrast to the WCR, one may derive the large sample properties for this estimator making formal inference possible. The estimator also allows for additional covariate effects.

Keywords Average hazard ratio · Cox model · Weighted Cox regression

1 Introduction

The Cox regression model is by far the most frequently used model to analyze survival data. The proportional hazards assumption gives a convenient way of reporting the effect of explanatory variables in terms of relative risks. This is however only appropriate if in fact the proportional hazards assumption is reasonable. An alternative to the Cox model is the Aalen additive hazards model Aalen (1980) that is well suited to

T. Martinussen (✉) · C. B. Pipper
Department of Biostatistics, University of Copenhagen, Øster Farimagsgade 5B,
1014 Copenhagen K, Denmark
e-mail: tma@sund.ku.dk

describe time-varying covariate effects. Reporting of results will inevitably be more complicated in this setting, however, and it is desirable with a simple quantity that summarizes the effect of a given covariate in a convenient way under such scenarios. An appealing measure is the odds-of-concordance, OC, defined as

$$\text{OC} = \frac{P(T^0 < T^1)}{1 - P(T^0 < T^1)}$$

imagining here a situation with two treatments, denoted 0 and 1, and with T^0 and T^1 being random variables describing the corresponding life times. In a recent paper, Schemper et al. (2009) advocated the use of weighted Cox regression (WCR) (Lin 1991; Sasieni 1993) as a general way of estimating an average hazard ratio (AHR) which may approximate the OC well under certain conditions. Their recommendation was based on comparing population values of AHR and OC under various conditions, and by extensive simulation studies, but no formal proof of properties was given. Estimation of the odds of concordance with censored data has also been considered in Dunkler et al. (2010) and Koziol and Jia (2009).

We suggest an alternative estimator based on the Aalen additive hazards model. It estimates a restricted version of the OC focussing on the time span where we observe death times. Its advantages are twofold. Firstly, it is easy to adjust for additional covariates, and secondly, its large sample properties can be derived. Specifically, we show that our estimator is consistent and asymptotically normal and we also provide a consistent estimator of the limiting variance. This estimator is calculated for a given covariate configuration. We also suggest an alternative estimator by averaging over the observed covariate distribution. It has the advantage that it does not depend on a given covariate value. This estimator is also shown to be consistent and asymptotically normal, and a consistent estimator of the limiting variance is provided. Small sample properties of the different estimators are investigated in a simulation study. Data from the TRACE study Jensen et al. (1997) concerning the effect of ventricular fibrillation on mortality for patients with myocardial infarction are analysed in Sect. 5.2—these data provide an example where the proportional hazards assumption is clearly inappropriate. In the Appendix we give proofs of the large sample properties of the proposed estimators. We also look into the properties of the WCR estimator (Appendix A.2 and A.3).

2 Odds of concordance and weighted Cox regression revisited

In this section we introduce some notation and review the recent proposal by Schemper et al. (2009) on how to estimate the OC by means of WCR (Lin 1991; Sasieni 1993). In the two sample case with only a single risk factor G present (the exposure) we show that this is a sensible approach when the effect of G is moderate and when the censoring does not depend on the exposure. Based on an extensive simulation study this was also the conclusion reached in Schemper et al. (2009). In Sect. 6 we comment on the exposure dependent censoring case. Let G denote the exposure variable and put $p = P(G = 1)$. The survival and hazard functions are denoted by S_g and h_g,

$g = 0, 1$. The population survival function, S, is thus given by $S = pS_1 + (1-p)S_0$. With this notation, and considering the case with no censorings, the estimator $\hat{\beta}_w$, from a WCR with weight function w converges in probability to the solution of

$$u_w(\beta) = \int_0^\infty w(t)\{e(t) - e(t,\beta)\}s^{(0)}(t)dt = 0, \tag{1}$$

see Lin (1991), where

$$e(t) = \frac{pS_1(t)h_1(t)}{(1-p)S_0(t)h_0(t) + pS_1(t)h_1(t)}, \quad e(t,\beta) = \frac{pS_1(t)\exp(\beta)}{(1-p)S_0(t) + pS_1(t)\exp(\beta)},$$
$$s^{(0)}(t) = (1-p)S_0(t)h_0(t) + pS_1(t)h_1(t).$$

The suggestion by Schemper et al. (2009) is to use the weight S thus resulting in the estimator $\hat{\beta}_S$. In practice, S needs of course to be replaced by the Kaplan-Meier estimator, $\hat{S}$. With this weight, and denoting $P(T^0 > T^1)$ by θ, straightforward calculations yield

$$u_S(\log(\mathrm{OC})) = p(1-p)\theta + p^2/2 - p\theta \int_0^\infty \frac{S(t)S_1(t)s^{(0)}(t)}{(1-p)(1-\theta)S_0(t) + p\theta S_1(t)}dt. \tag{2}$$

Notice that for $\log(\mathrm{OC}) = 0$, that is, $\theta = 1/2$, the last term on the right hand side of (2) simplifies to

$$p\int_0^\infty S_1(t)s^{(0)}(t)dt = p(1-p)/2 + p^2/2 = p/2$$

from which we see that the right hand side of (2) is zero. Hence, in this case, the estimator from the WCR indeed converges in probability to $\log(\mathrm{OC})$. Intuitively, the derivations above lead us to expect that this estimator will be close to $\log(\mathrm{OC})$ when the latter quantity is close to zero, that is when the effect of G is moderate. In general, however, one may not expect this to be the case. We explore this in more detail in the Appendix where we study a scenario constructed to mimic the application of this paper, that is, with a large effect of the exposure for an initial period of time, see Sect. 5.2. In the next section we suggest an alternative estimator of the log odds concordance that also allows for additional covariates. Furthermore, large sample properties for this estimator of the OC can be derived making formal inference possible.

3 Model and definition of odds of concordance

Let T_i denote the possibly right-censored failure time of the ith individual. Furthermore suppose that for the ith individual we have recorded the covariates G_i, X_i, where G_i is the exposure variable corresponding to no treatment or treatment and X_i is a p-dimensional vector of additional covariates. With this setup we assume the Aalen additive hazards model (Aalen 1980)

$$\lambda(t, G_i = g, X_i = x) = \beta_0(t) + \beta_G(t)g + x^T \beta_X(t),$$

where $\beta_0(t)$ is the baseline hazard function, $\beta_G(t)$ is the excess hazard due to treatment, and $\beta_X(t)$ denotes the effect of the additional covariates. To define the odds of concordance within this model suppose that T^0 and T^1 are two independent failure times that follow the above model with $g = 0$ and $g = 1$, respectively, and with the same value, x, of the additional covariates. With this notation, the odds of concordance is given by

$$\mathrm{OC}(x) = \frac{P_x(T^0 > T^1)}{P_x(T^1 > T^0)}, \tag{3}$$

with

$$P_x(T^1 > T^0) = \int_0^\infty \exp[-\{2B_0(t) + B_G(t) + 2x^T B_X(t)\}]\{dB_0(t) + x^T dB_X(t)\},$$

where $B_0(t) = \int_0^t \beta_0(s)ds$ and similarly with the other quantities. Notice that with $x = 0$ the above expression reduces to the average hazard ratio as originally defined in Efron (1967) and later extended by Kalbfleisch and Prentice (1981). In both these papers modification of the average hazard ratio to ensure stability of the suggested estimates in the presence of censored event times was necessary. We consider the following modification

$$\mathrm{OC}(x, v) = \frac{P_x(T^1 < T^0, T^1 < v)}{P_x(T^0 < T^1, T^0 < v)}, \tag{4}$$

where we suggest to choose v as $v = \psi^{-1}(q)$, $q \in [0, 1]$, with

$$\psi(t) = P_0(\max(T^0, T^1) > t).$$

The parameter $q \in [0, 1]$ relates to the quantiles in the distribution corresponding to the survival function given by ψ. Thus, we may choose q irrespective of the actual distribution of the data and as shown later still obtain stability of the estimators we suggest. This is contrary to the data specific suggestion given by for instance Kalbfleisch and Prentice (1981). The $\mathrm{OC}(x, v)$ is a restricted version of the OC-measure focussing on the time frame where we actually have data. In some applications the choice of v may be guided by substance matter if one wishes to consider some pre specified time interval. With $q = 0$, (4) reduces to (3). In the case without any additional covariates we will adopt the notation $\mathrm{OC}(v)$ for (4).

Before suggesting an estimator of $\mathrm{OC}(x, v)$ we express it as a function of the cumulated regression functions of the Aalen model. With $A_0 = (1, 0, x^T)$, $A_1 = (1, 1, x^T)$, $A_2 = (2, 1, 2x^T)$, $B(t) = \{B_0(t), B_G(t), B_X^T(t)\}^T$, we can write

$$P_x(T^k < T^l, T^k < v) = \int_0^v e^{-A_2 B(u)} A_k dB(u)$$

with $(k, l) = (1, 0)$ and $(k, l) = (0, 1)$. The quantity $\mathrm{OC}(x, v)$ is a measure of the exposure effect for a given covariate configuration, x. It may be attractive to report a single number that does not depend on a specific covariate configuration, which is also the case for the WCR measure. Here we suggest an estimator based on the marginal probabilities

$$p_1(v) = \int P_x(T^1 < T^0, T^1 < v) dF(x),$$
$$p_0(v) = \int P_x(T^0 < T^1, T^0 < v) dF(x),$$

where dF denotes integration w.r.t. the distribution of the covariates X_i. We therefore suggest to estimate the quantity

$$\mathrm{OC}(v) = \frac{p_1(v)}{p_0(v)}, \tag{5}$$

which we shall term the overall adjusted OC. In the next section, we use the empirical counterparts of $p_0(v)$ and $p_1(v)$ to obtain an estimator of this latter quantity.

4 Estimation and large sample properties

Suppose that the event times T_i are subject to right-censoring, that is they are not observed if they exceed the censoring time C_i. Instead we observe the first time either failure or censoring occurs $\tilde{T}_i = \min(T_i, C_i)$ and an indicator of whether it is censoring or failure that occurs $\delta_i = I(T_i \leq C_i)$. We shall assume that $\{T_i, C_i, G_i, X_i\}_i$ are independent and identically distributed according to the model described in the previous section, and that T_i and C_i are conditionally independent given (G_i, X_i). This is the standard "independent censoring" assumption that allows for exposure dependent censoring; we discuss this further in Sect. 6 in relation to the WCR estimator. These quantities translate into the counting process framework of Andersen et al. (1993) as the counting process $N_i(t) = \delta_i I(\tilde{T}_i \leq t)$ and at risk process $Y_i(t) = I(\tilde{T}_i \geq t)$ for the ith individual. With this notation, the estimator of $B(t)$ is obtained by well-established methods Martinussen and Scheike (2006) as

$$\hat{B}(t) = \int_0^t Z^-(s) dN(s), \tag{6}$$

where $Z^-(t)$ is the generalized inverse of $Z(t)$ with the latter being the $n \times (p + 2)$-matrix with ith row $Y_i(t)(1, G_i, X_i^T)$.

Estimators $\hat{a}(x, v)$ and $\hat{b}(x, v)$ of the probabilities $P_x(T^1 < T^0, T^1 < v)$ and $P_x(T^0 < T^1, T^0 < v)$ are obtained by simply plugging in the above estimator of B. These estimators are then used to estimate $\mathrm{OC}(x, v)$ defined in (4) as

$$\hat{\mathrm{OC}}(x, v) = \frac{\hat{a}(x, v)}{\hat{b}(x, v)}.$$

Furthermore we estimate the marginal probabilities $p_1(v)$ and $p_0(v)$ by

$$\hat{p}_1(v) = n^{-1} \sum_{i=1}^{n} \hat{a}(X_i, v), \quad \hat{p}_0(v) = n^{-1} \sum_{i=1}^{n} \hat{b}(X_i, v),$$

respectively. From this we may estimate $\mathrm{OC}(v)$ defined in (5) by

$$\hat{\mathrm{OC}}(v) = \frac{\hat{p}_1(v)}{\hat{p}_0(v)}. \tag{7}$$

Since we have the following asymptotic representation

$$n^{1/2}\{\hat{B}(t) - B(t)\} = n^{-1/2} \sum_{i=1}^{n} \epsilon_i^B(t) + o_p(1), \tag{8}$$

where the ϵ_i^B's are zero-mean iid terms, see Martinussen and Scheike (2006, Chap. 5), we may show that

$$n^{1/2}\{\hat{\mathrm{OC}}(x, v) - \mathrm{OC}(x, v)\} = n^{-1/2} \sum_{i=1}^{n} \epsilon_i^{\mathrm{OC}}(x, v) + o_p(1),$$

with $\{\epsilon_i^{\mathrm{OC}}(x, v)\}$ being zero-mean iid terms. Hence, $n^{1/2}\{\hat{\mathrm{OC}}(x, v) - \mathrm{OC}(x, v)\}$ converges in distribution to a zero-mean normal variate with a variance that is consistently estimated by

$$n^{-1} \sum_{i=1}^{n} \hat{\epsilon}_i^{\mathrm{OC}}(x, v)^2.$$

The proof of this result is given in the Appendix A.1, where we also give an expression for $\hat{\epsilon}_i^{\mathrm{OC}}(x, v)$. We also show in the Appendix that

$$n^{1/2}\{\hat{p}_1(v) - p_1(v)\} = n^{-1/2} \sum_{i=1}^{n} \epsilon_i^a(v) + o_P(1), \tag{9}$$

$$n^{1/2}\{\hat{p}_0(v) - p_0(v)\} = n^{-1/2} \sum_{i=1}^{n} \epsilon_i^b(v) + o_P(1), \tag{10}$$

from which asymptotic normality of $\hat{\mathrm{OC}}(v)$ is deduced and a consistent estimator of the asymptotic variance is obtained as above. Expressions for $\epsilon_i^a(v)$ and $\epsilon_i^b(v)$ are given in Appendix A.1.

5 Numerical results

5.1 Simulation study

In the simulation study we mimic the TRACE data considered in the next subsection by utilizing estimated parameter values based on the model

$$h_1(t) = \gamma h_0(t) I(t \leq r) + h_0(t) I(t > r)$$

with a piecewise constant baseline hazard that also changes in the change point r. Specifically the change point for the TRACE data was estimated in Martinussen and Scheike (2007) to be 0.092 years. Furthermore the baseline hazard was estimated to be

$$h_0(t) = 0.8 \cdot I(t < 0.092) + 0.09 \cdot I(t \geq 0.092)$$

and the effect γ of the exposure in the initial period was estimated to exp(2.05). We generated the exposure from a Bernoulli distribution with probability p_{sim} given by $p,\ 2p,\ 4p$, where $p = 71/1000$ is the frequency of exposure in the TRACE data. Event times are sampled from the above specifications with the effect size γ_{sim} given by $\gamma = \exp(2.05)$. Number of observations was taken to be $n = 200, 600, 1000$. We induced censoring times similarly to those seen in the TRACE study, which are uniformly distributed between 6 and 8 years resulting in approximately 50 % censoring. For each configuration of n and p_{sim}, 2000 data sets were generated and for each data set $\log\{OC(v)\}$ as defined in (4) is estimated by the proposed method for $q = 0.9, 0.7$ with $v = \psi^{-1}(q)$. Results are given in Table 1.

Table 1 Summary results of the first simulation study. Bias corresponds to sample median minus true value (on log-scale)

p_{sim}	n	q=0.9				q=0.7			
		bias	SE	SEE	CP	bias	SE	SEE	CP
p	200	−0.10	0.49	0.44	94.1	−0.11	0.44	0.39	93.3
$2p$		−0.03	0.355	0.34	94.9	−0.06	0.31	0.30	95.2
$4p$		0.00	0.29	0.29	95.0	−0.02	0.23	0.24	95.6
p	600	−0.04	0.27	0.27	95.1	−0.03	0.25	0.24	94.6
$2p$		−0.02	0.20	0.20	95.0	−0.02	0.19	0.18	94.8
$4p$		−0.01	0.17	0.17	94.9	−0.01	0.14	0.14	95.2
p	1000	−0.02	0.21	0.21	95.4	−0.03	0.19	0.19	94.4
$2p$		0.00	0.16	0.16	95.1	−0.01	0.14	0.14	94.9
$4p$		0.00	0.13	0.13	95.1	−0.01	0.11	0.11	95.5

SE corresponds to sample standard error of the log $(OC(q))$ estimates and SEE denotes the median of the estimated standard errors. CP denotes the 95 % coverage probability. The parameter p is the frequency of exposure seen in the TRACE study that is $p = 0.07$, and q defines the $(1 - q)$-quantile of the distribution $\max(T^0, T^1)$

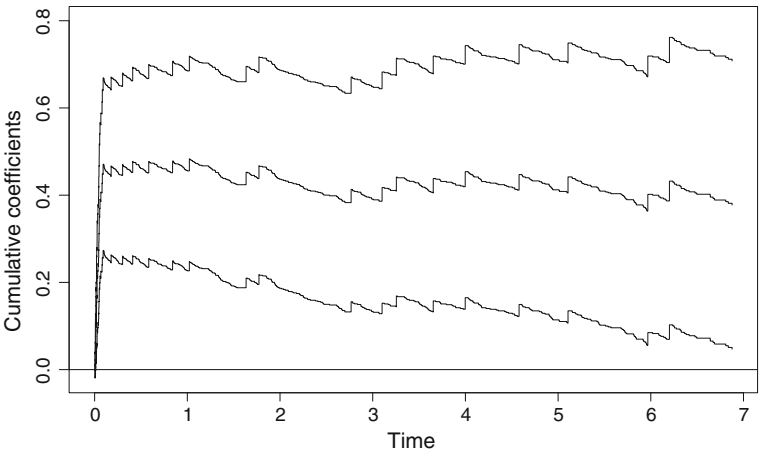

Fig. 1 TRACE data. Effect of ventricular fibrilation shown as estimate of cumulative regression coefficient along with 95 % confidence limits

For the low sample size ($n = 200$) the method proposed in this paper shows some bias when the frequency p_{sim} is at its lowest. This is not surprising since in this case the expected number of individuals in the exposure group is only 14. In all other simulation scenarios the performance of the proposed procedure in terms of coverage probabilities is good with coverage probabilities close to the nominal 95 %.

We also conducted a small simulation study to investigate the small sample properties of the overall adjusted OC estimator defined in (7). The same setup as in the previous study was considered but only with the configuration $p_{sim} = 4p$ and $n = 200$. In addition we include an exponentially distributed covariate in the hazard with a corresponding regression coefficient equal to 0.5. The true value of the overall adjusted OC for $q = 0.7$ in this scenario is calculated to 2.14. The sample median of the estimator on log-scale minus the true log-value is -0.02 with a sample standard error of 0.24. The median estimated standard error is 0.22 and the 95 % coverage probability 0.94. Based on this we conclude that the performance of the estimation procedure is acceptable.

5.2 The TRACE study

The TRACE study group Jensen et al. (1997) investigated the prognostic importance of various risk factors on mortality for approximately 6600 patients with myocardial infarction. We consider a random subsample of 1000 of these patients that have previously been analysed in Martinussen and Scheike (2006, Chap. 5) and Martinussen and Scheike (2007); the data are available in the R-package `timereg`. In these analyses, ventricular fibrillation was recorded for 71 patients and was identified to be a very important risk factor. Figure 1 displays the estimated cumulative regression coefficient for this variable fitted in the additive Aalen model where we also adjust for gender, diabetes (present/absent), clinical heart failure status (present/absent), and age

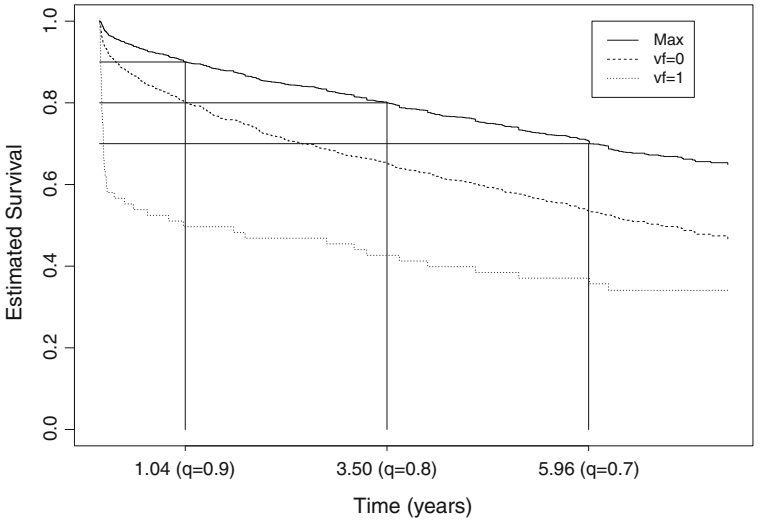

Fig. 2 TRACE data. Estimated survival and q values

Table 2 Results from the analysis of TRACE data

Model	$\hat{OC}(1.04)$ (95 % CI)	$\hat{OC}(3.5)$ (95 % CI)	$\hat{OC}(5.96)$ (95 % CI)
Un–adjusted	3.76 (2.54–5.57)	2.65 (1.81–3.89)	2.28 (1.57–3.31)
Conditional	5.61 (3.08–10.2)	3.54 (2.15–5.84)	2.71 (1.71–4.30)
Overall adjusted	4.06 (2.77–5.95)	2.94 (2.00–4.31)	2.60 (1.76–3.83)

The conditional estimate corresponds to covariate configuration x given by gender = females, diabetes status = absent, clinical heart failure = absent and age at enrollment equal to 70 years. The OC is evaluated at time points $v = 1.04, 3.5, 5.96$ corresponding to $q = 0.9, 0.8, 0.7$, respectively, with $v = \psi^{-1}(q)$

at enrollment centered around 70 years. A very strong effect is seen, but it vanishes after approximately one month. Specifically, Martinussen and Scheike (2007), using a change-point model, estimated the change-point to 33 days. In the study design individuals were followed for approximately 6 years after which their follow-up was terminated according to a uniformly distributed censoring time in the range 6–8 years.

During follow-up, 47 patients in the vf = 1 group and 466 patients in the vf = 0 group died from myocardial infarction. The estimated survival curves for each of the two groups as well as an estimate of ψ (denoted Max) are depicted in Fig. 2.

First we consider estimating (4) only including ventricular fibrillation in the model. From the estimate of ψ depicted in Fig. 2 it is evident that we may estimate (4) for $q \geq 0.7$. Estimates of (4) for $q = 0.9, 0.8, 0.7$ (corresponding roughly to 1, 3.5 and 6 years) are obtained along with 95 % confidence limits calculated on a log-scale and back transformed. Similar conditional estimates are calculated for the covariate configuration x: gender = Females, Diabetes status = absent, clinical heart failure = absent, age at enrollment= 70 years, corresponding to the baseline intensity. Finally, we calculate overall adjusted OC estimates given by (7) for $q = 0.9, 0.8, 0.7$. Table 2 contains the results of these analyses.

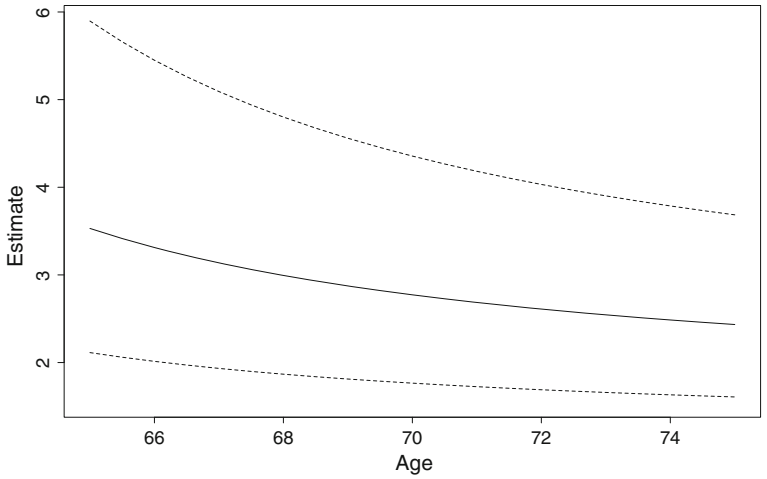

Fig. 3 TRACE data. $OC(x, 6)$ for different ages

All estimates indicate a substantial effect of ventricular fibrillation. Adjusting for the other explanatory variables yields conditional estimates of $OC(x, v)$ that are higher than the unadjusted estimates for the risk factor configuration we consider which corresponds to a person of median age with a low risk profile in the sampled population w.r.t. the other considered risk factors. As noted earlier the $OC(x, v)$ depends on the covariate configuration x. This is illustrated in Fig. 3 that displays $OC(x, v)$ with v equal to 6 years ($q \approx 0.7$) in the age span 65–75 years keeping the other covariates fixed at: gender = Females, Diabetes = absent and clinical heart failure = absent. We finally calculate the overall adjusted OC given by (7); this measure does not depend on a specific covariate configuration similarly to the WCR estimate. The adjusted overall OC estimates are seen to be lower than the specific conditional OC estimates, where we condition on a low risk profile on the other risk factors (sex, clinical heart failure and diabetes). The adjusted WCR estimator is of similar magnitude and precision as the overall adjusted OC estimates.

6 Concluding remarks

The WCR estimator is attractive as it summarizes the exposure effect by one number. We suggested a similar simple summary of the exposure effect based on the marginal concordance probabilities obtained by averaging over the empirical covariate distribution. This estimator was shown to be asymptotically normal and a consistent estimator of the variance was provided justifying formal inference, something which is problematic for the WCR estimator due to its unknown large-sample properties. In Appendix A.2 we study the potential bias of the WCR closer under a scenario similar to that of the TRACE study, but without censoring. We give a theoretical result showing that WCR may be biased under very strong exposure effect; this is further investigated in some simulations. They point to that the bias seems to be of little

practical consequence. However, it may be more problematic to employ the WCR estimator when the censoring depends on the exposure (still assuming that T and C are conditionally independent given covariates). This is explored in Appendix A.3 where it is seen that WCR may indeed be biased.

The overall adjusted $OC(v)$ given in (5) is constructed by plugging in the marginal concordance probabilities. These are obtained from the full model by integrating over the conditional concordance probabilities with respect to the covariate distribution. An alternative more direct approach would have been to average the conditional $OC(x, v)$ or $\log\{OC(x, v)\}$ with respect to the covariate distribution. However, by the above approach we can still interpret (5) in terms of a comparison between two randomly chosen exposed and unexposed individuals that are otherwise comparable (alike with respect to the other risk factors). Interpretation of the alternative approaches if $OC(x, v)$ varies with x is unclear.

Appendix

A.1 Large sample properties of proposed estimators

By straightforward calculations one has

$$n^{1/2}\{\hat{a}(x, v) - a(x, v)\} = \int_0^v e^{-A_2 B(t)} A_1 dn^{1/2}\{\hat{B} - B)(t) \\ - \int_0^v e^{-A_2 B(t)} A_2 n^{1/2}\{\hat{B}(t) - B(t)\} A_1 dB(t) + o_p(1)$$

and because of (8) it is easy to see that we can write

$$n^{1/2}\{\hat{a}(x, v) - a(x, v)\} = \sum_{i=1}^{n} \epsilon_i^a(x, v) + o_p(1),$$

where ϵ_i^a are zero-mean iid variates given by

$$\epsilon_i^a(x, v) = \int_0^v e^{-A_2 B(t)} A_1 d\epsilon_i^B(t) - \int_0^v e^{-A_2 B(t)} A_2 \epsilon_i^B(t) A_1 dB(t).$$

Similarly one can derive that

$$n^{1/2}\{\hat{b}(x, v) - b(x, v)\} = \sum_{i=1}^{n} \epsilon_i^b(x, v) + o_p(1),$$

with

$$\epsilon_i^b(x, v) = \int_0^v e^{-A_2 B(t)} A_0 d\epsilon_i^B(t) - \int_0^v e^{-A_2 B(t)} A_2 \epsilon_i^B(t) A_0 dB(t).$$

The result now follows by noting that

$$n^{1/2}\{\hat{\mathrm{OC}}(x,v)-\mathrm{OC}(x,v)\}=$$
$$\left[n^{1/2}\{\hat{a}(x,v)-a(x,v)\}-n^{1/2}\{\hat{b}(x,v)-b(x,v)\}\mathrm{OC}(x,v)\right]/b(x,v)+o_p(1)$$

The expression for $\hat{\epsilon}_i^{\mathrm{OC}}(x,v)$ is given by

$$\frac{1}{\hat{b}(x,v)}\{\hat{\epsilon}_i^a(x,v)-\hat{\mathrm{OC}}(x,v)\hat{\epsilon}_i^b(x,v)\},$$

where $\hat{\epsilon}_i^a(x,v)$ is obtained from $\epsilon_i^a(x,v)$ by replacing unknown quantities with their empirical counterparts, similarly with $\hat{\epsilon}_i^b(x,v)$; for $\hat{\epsilon}_i^B(t)$, see Martinussen and Scheike (2006).

We now turn to the asymptotic distribution of $\hat{p}_1(v)$ and $\hat{p}_0(v)$. Following the line of arguments in Chen et al. (2010, Appendix A) one may show that the $\epsilon_i^a(v)$ and $\epsilon_i^b(v)$ in display (9) and (10) are given by

$$\epsilon_i^a(v)=a(X_i,v)-p_1(v)+E_X\epsilon_i^a(X,v)\}$$
$$\epsilon_i^b(v)=b(X_i,v)-p_0(v)+E_X\epsilon_i^b(X,v),$$

where E_X denotes expectation w.r.t. the distribution of X. A further calculation shows that

$$E_X\epsilon_i^a(X,v)=\int_0^v E_X f_1(t,X,B)d\epsilon_i^B(t)-\int_0^v E_X f_2(t,X,B,\epsilon_i^B)dB(t),$$
$$E_X\epsilon_i^b(X,v)=\int_0^v E_X f_3(t,X,B)d\epsilon_i^B(t)-\int_0^v E_X f_4(t,X,B,\epsilon_i^B)dB(t),$$

with

$$f_1(t,x,B)=e^{-A_2B(t)}A_1,\quad f_2(t,x,B,\epsilon_i^B)=e^{-A_2B(t)}A_2\epsilon_i^B(t)A_1,$$
$$f_3(t,x,B)=e^{-A_2B(t)}A_0,\quad f_4(t,x,B,\epsilon_i^B)=e^{-A_2B(t)}A_2\epsilon_i^B(t)A_0.$$

The above may then be combined into expressions for $\epsilon_i^a(v)$ and $\epsilon_i^b(v)$ the empirical counterpart of which may be obtained by inserting the estimator of B, the empirical counterpart of $\epsilon_i^B(t)$, and replacing expectations by their empirical counterparts. Asymptotic normality around the true value for $\hat{\mathrm{OC}}(v)$ now follows from

$$n^{1/2}\{\hat{\mathrm{OC}}(v)-\mathrm{OC}(v)\}=n^{1/2}\{\hat{p}_1(v)-p_1(v)\}/p_0(v)-\mathrm{OC}(v)n^{1/2}\{\hat{p}_0(v)-p_0(v)\}$$

that also gives the asymptotic variance.

A.2 Bias of the WCR estimator: the uncensored case

As noted in Sect. 2 the WCR estimator converges in probability to the parameter, β^* say, that solves (1). If OC=1 it was seen that $e^{\beta^*}=\mathrm{OC}$. In general, however, one

may not expect this to be the case. We will explore this in some more detail under a scenario that mimics the application given in Sect. 5.2, that is, with a large effect of the exposure for an initial period of time. Specifically, we consider the situation where

$$h_1(t) = \gamma h_0(t) I(t \leq r) + h_0(t) I(t > r) \tag{11}$$

corresponding to an extended Cox-model with relative risk γ in the time period $[0, r]$ and with relative risk 1 thereafter. This results in the following relationship between the two survival functions

$$S_1(t) = S_0(t)^{\gamma} I(t \leq r) + S_0(r)^{\gamma - 1} S_0(t) I(t > r). \tag{12}$$

In this case a straightforward calculation reveals that

$$\theta = \frac{\gamma}{1+\gamma} + S_0(r)^{\gamma+1}(\frac{1}{2} - \frac{\gamma}{1+\gamma})$$

from which we get

$$\mathrm{OC} = \frac{\gamma + (\gamma+1) S_0(r)^{\gamma+1}(\frac{1}{2} - \frac{\gamma}{1+\gamma})}{1 - (\gamma+1) S_0(r)^{\gamma+1}(\frac{1}{2} - \frac{\gamma}{1+\gamma})}. \tag{13}$$

Note from this expression that $\mathrm{OC} \leq \gamma$ for $\gamma > 1$ as we would expect since the exposure effect is not present after time r. The following result gives a bound on the bias.

Proposition 1 *Under model (11) and with $\gamma \to \infty$ and $r \to 0$ such that $\gamma S_0(r)^{\gamma} \to c$ for some constant $c > 0$ then, for $n \to \infty$, $\hat{\beta}_S$ converges in probabilty to β^* with*

$$\log(OC) - \beta^* \geq \frac{p(1-p)^2}{2\{(1-p)(1/c + 1/2) + p\}}. \tag{14}$$

Proof We first show that $u_S\{\log(\mathrm{OC})\}$ converges to a constant different from zero. Under model (12) we may rewrite (2) as

$$\begin{aligned} & u_S\{\log(\mathrm{OC})\} = p(1-p)\theta + p^2/2 \\ & -p\mathrm{OC} \int_0^r \frac{\{(1-p)S_0(t) + pS_0(t)^{\gamma}\} S_0(t)^{\gamma} \{(1-p)S_0(t) + \gamma p S_0(t)^{\gamma}\} \lambda_0(t)}{(1-p)S_0(t) + \mathrm{OC} p S_0(t)^{\gamma}} dt \\ & -p\theta \int_r^{\infty} \frac{\{1 - p + pS_0(r)^{\gamma-1}\}^2 S_0(r)^{\gamma-1}}{(1-p)(1-\theta) + p\theta S_0(r)^{\gamma-1}} S_0(t)^2 \lambda_0(t) dt. \end{aligned}$$

Also note that from the expression (13) of OC we get $\mathrm{OC}/\gamma \to \frac{1}{1+c/2}$. By the substitution $u = S_0(t)^{\gamma}$ we see that the second term on the right hand side in the above equality may be re-expressed as

$$p\frac{\mathrm{OC}}{\gamma}\int_{S_0(r)^\gamma}^{1}\frac{\{(1-p)u^{1/\gamma}+pu\}\{(1-p)u^{1/\gamma}+\gamma pu\}}{(1-p)u^{1/\gamma}+\mathrm{OC}pu}du.$$

For any $u \in (S_0(r)^\gamma, 1)$ we have that under the convergence specified above

$$\frac{\{(1-p)u^{1/\gamma}+pu\}\{(1-p)u^{1/\gamma}+\gamma pu\}}{(1-p)u^{1/\gamma}+\mathrm{OC}pu} \to (1-p+pu)(1+c/2).$$

Furthermore for $\gamma > 1$

$$\sup_{u\in(S_0(r)^\gamma,1)}\frac{\{(1-p)u^{1/\gamma}+pu\}\{(1-p)u^{1/\gamma}+\gamma pu\}}{(1-p)u^{1/\gamma}+\mathrm{OC}pu} \le \frac{1-p+\gamma p}{1-p+\mathrm{OC}p} \to 1+c/2.$$

Thus by Lebesgue's Dominated Convergence Theorem

$$p\frac{\mathrm{OC}}{\gamma}\int_{S_0(r)^\gamma}^{1}\frac{\{(1-p)u^{1/\gamma}+pu\}\{(1-p)u^{1/\gamma}+\gamma pu\}}{(1-p)u^{1/\gamma}+\mathrm{OC}pu}du \to p(1-p)+p^2/2.$$

The third term on the right hand side in the above equality may be directly calculated as

$$\frac{p(1-p+pS_0(r)^{\gamma-1})^2}{(1-p)\mathrm{OC}^{-1}S_0(r)^{1-\gamma}+p}S_0(r)^2/2.$$

Hence, under the considered convergence, $u_S\{\log(\mathrm{OC})\}$ converges to

$$-\frac{p(1-p)^2}{2\{(1-p)(1/c+1/2)+p\}}.$$

By similar arguments it may be seen that

$$u_S^{'}\{\beta\} \to -\int_0^1 \frac{\{(1-p)+pu\}pu}{e^\beta p^2u^2+2(1-p)pu+(1-p)^2e^{-\beta}}\,du \ge -1.$$

for $\beta > 0$. By the mean-value theorem we then have

$$\log(\mathrm{OC})-\beta_0 = u_S\{\log(\mathrm{OC})\}/u_S^{'}\{\beta^*\},$$

where β^* is on the line segment between $\log(\mathrm{OC})$ and β_0. The result thus follows from the latter display.

We now explore the magnitude of the potential bias of the WCR in scenarios similar to the TRACE data. We use estimated parameter values based on the model (11) with a piecewise constant baseline hazard that also changes in the change point r. Specifically the change point for the TRACE data was estimated in Martinussen and

Table 3 Bias of the WCR estimator, corresponding to sample median minus the true value

(p_{sim}, γ_{sim})	$p_{sim} = p$			$p_{sim} = 2p$			$p_{sim} = 4p$		
	γ	2γ	4γ	γ	2γ	4γ	γ	2γ	4γ
OC	2.16	4.26	12.3	2.16	4.26	12.3	2.16	4.26	12.3
Bias	−0.01	−0.03	−0.13	0.00	−0.06	−0.22	−0.02	−0.10	−0.33
Bound	−0.04	−0.04	−0.04	−0.07	−0.07	−0.06	−0.09	−0.09	−0.08

Bound corresponds to the theoretical bound given in Proposition 1. The parameter p and γ are the frequency of exposure and the estimated exposure effect (see text for more details) in the TRACE study, that is, $p = 0.07$ and $\gamma = \exp(2.05)$, respectively

Scheike (2007) to be 0.092 years. Furthermore the baseline hazard was estimated to be

$$h_0(t) = 0.8 \cdot I(t < 0.092) + 0.09 \cdot I(t \geq 0.092)$$

and the effect γ of the exposure in the initial period, see (11), was estimated to exp(2.05). Finally the frequency of exposure was also taken from the TRACE data, that is, $p = 71/1000$. We simulate data sets consisting of 1000 binary exposures sampled from a Bernoulli distribution with probability p_{sim} given by p, $2p$, $4p$ and 1000 event times sampled from the above specifications with the effect size γ_{sim} given by γ, 2γ, 4γ. For each combination of p_{sim} and γ_{sim} 2000 data sets were generated. For each data set $\hat{\log}\{OC\}$ is estimated by means of weighted Cox regression as suggested in Schemper et al. (2009). The median bias (on log-scale) is reported in the below Table 3.

The results in Table 3 confirm the claim of Proposition 1 as an apparent bias of WCR is seen for the scenarios with large p_{sim} and γ_{sim}. Also notice that for $\gamma_{sim} = 4\gamma$ the bias is substantial and well above the bound given in Proposition 1. However, an effect of 4γ is unrealistic in practice and we therefore conclude that the bias of WCR does not seem to pose a practical problem in this setting when there is no censoring taking place.

A.3 Bias of the WCR estimator: censoring depending on exposure

The performance of the WCR under censoring not depending on the exposure (or any other important covariate) seems to be similar to what we reported in the previous subsection; this is also in line with what was reported by Schemper et al. (2009). Note that the weight under censoring should be $\hat{S}(t)/\hat{H}(t)$, where $\hat{H}(t)$denote the Kaplan-Meier estimator of the censoring distribution Schemper et al. (2009). We will now investigate an exposure dependent censoring situation. Suppose that censoring times are distributed according to survival functions H_0 and H_1 in the two exposure groups. Then the characterization of the WCR estimator is in terms of the solution to

$$u_w(\beta) = \int_0^\infty w(t)\{e(t) - e(t, \beta)\}s^{(0)}(t)dt = 0,$$

Table 4 Comparison of bias of the WCR and proposed estimators on the log scale

		WCR		Proposal	
n	Cens. %	*Bias*	95 % CP	*Bias*	95 % CP
500	0.80	−0.30	0.85	−0.17	0.89
	0.50	0.02	0.94	−0.02	0.94
	0.20	0.00	0.95	0.01	0.95
1000	0.80	−0.32	0.77	−0.11	0.93
	0.50	0.02	0.95	−0.01	0.93
	0.20	0.00	0.96	0.00	0.95
2000	0.80	−0.32	0.57	−0.05	0.93
	0.50	0.02	0.95	−0.01	0.95
	0.20	−0.00	0.95	−0.00	0.94

Bias corresponds to sample median of the WCR estimator minus the true value. Also reported is the sample coverage percentage by means of 95 % Wald CI calculated for the logarithm of the WCR estimator. A similar summary is made for the proposed estimator

where

$$
\begin{aligned}
w(t) &= \{(1-p)S_0(t) + pS_1(t)\}\{(1-p)H_0(t) + pH_1(t)\}^{-1},\\
s^{(0)}(t) &= (1-p)S_0(t)H_0(t)h_0(t) + pS_1(t)H_1(t)h_1(t),\\
e(t) &= \frac{pS_1(t)H_1(t)h_1(t)}{(1-p)S_0(t)H_0(t)h_0(t) + pS_1(t)H_1(t)h_1(t)},\\
e(t,\beta) &= \frac{pS_1(t)H_1(t)\exp(\beta)}{(1-p)S_0(t)H_0(t) + pS_1(t)H_1(t)\exp(\beta)}.
\end{aligned}
$$

Clearly, in this scenario, the censoring distribution does not cancel out as would be the case if $H_0 = H_1$. As a consequence the solution to $u_w(\beta) = 0$ may depend on censoring in which case the solution is clearly not given by $\log\{OC\}$. To investigate this we simulated event times according to the hazard rate

$$h_g(t) = 0.5 \cdot (1 + g \cdot 1.86 \cdot t),$$

where $g = 0, 1$ denotes exposure groups. The censoring times in the exposed group ($g = 1$) are given by $I_i\tau_1 + (1 - I_i)\tau$, where the censoring indicator I_i at τ_1 is generated as Bernoulli variables with $P(I_i = 1) = p_{cens}$. In the unexposed group ($g = 0$) failure times were censored at τ. Finally, the G_i's are generated as Bernoulli variables with $P(G_i = 1) = P(G_i = 0) = 0.5$.

We consider the scenarios $n = 500, 1000, 2000$. For censoring, the scenarios $(p_{cens}, \tau_1) = (0.93, 0.25),\ (0.58, 0.25),\ (0.53, 1)$ are considered. The two first of these scenarios correspond to early censoring in the exposure group with 80 and 50 % censoring before τ. The third scenario corresponds to late censoring in the exposure group with 20 % censoring before τ. Finally, τ is set to 2.

For each scenario 2000 datasets were generated, and for each data set, the WCR estimator was computed as suggested in Schemper et al. (2009). Table 4 below shows the performance of the WCR estimator and the proposed estimator under these scenarios

Table 4 shows that the WCR estimator is unbiased with coverage close to the nominal 95 % when censoring in the exposure group is not large. However in the case of strong early censoring the WCR estimator is biased and this compromises the coverage. In comparison, for $n = 500$, the proposed estimator seems to be less biased but with a poor coverage when early censoring is strong. For $n = 1000$ and 2000 the bias decreases and coverages are satisfactory.

References

Aalen OO (1980) A model for non-parametric regression analysis of counting processes. In: Klonecki W, Kozek A, Rosinski J (eds) Lecture notes in statistics-2: mathematical statistics and probability theory. Springer, New York, pp 1–25

Andersen PK, Borgan Ø, Gill RD, Keiding N (1993) Statistical models based on counting processes. Springer, New York

Chen L, Lin DY, Zeng D (2010) .Biometrika 97:713–726

Dunkler D, Schemper M, Heinze G (2010) Gene selection in microarray survival studies under possibly non-proportional hazards. Bioinformatics 26:784–790

Efron B (1967) The two sample problem with censored data. Proc 5th Berkeley Symp 4:831–853

Jensen GV, Torp-Pedersen C, Hildebrandt P, Kober L, Nielsen FE, Melchior T, Joen T, Andersen PK (1997) Does in-hospital ventricular fibrillation affect prognosis after myocardial infarction?. Eur Heart J 18:919–924

Kalbfleisch JD, Prentice RL (1981) Estimation of the average hazard ratio. Biometrika 68:105–112

Koziol JA, Jia Z (2009) The concordance index C and the Mann-Whitney parameter Pr(X>Y) with randomly censored data. Biometr J 51:1521–4036

Lin D (1991) Goodness-of-fit analysis for the cox regression model based on a class of parameter estimators. J Am Stat Assoc 86:725–728

Martinussen T, Scheike TH (2006) Dynamic regression models for survival data. Springer, New York

Martinussen T, Scheike TH (2007) Aalen additive hazards change-point model. Biometrika 94:861–872

Sasieni P (1993) Weighted partial likelihood estimators for the cox model. J Am Stat Assoc 88:144–152

Schemper M, Wakounig S, Henize G (2009) Estimation of average hazard ratios by weighted Cox regression. Stat Med 28:2473–2489

Vaart AWvan der , Wellner JA (1996) Weak convergence and empirical processes. Springer, New York