Published in final edited form as:

Lifetime Data Anal. 2016 April; 22(2): 263–279. doi:10.1007/s10985-015-9330-3.

Estimating the Concordance Probability in a Survival Analysis with a Discrete Number of Risk Groups

Glenn Heller^{1,*} and Qianxing Mo²

¹Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, U.S.A.

²Department of Medicine and Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030, U.S.A.

Abstract

A clinical risk classification system is an important component of a treatment decision algorithm. A measure used to assess the strength of a risk classification system is discrimination, and when the outcome is survival time, the most commonly applied global measure of discrimination is the concordance probability. The concordance probability represents the pairwise probability of lower patient risk given longer survival time. The c-index and the concordance probability estimate have been used to estimate the concordance probability when patient-specific risk scores are continuous. In the current paper, the concordance probability estimate and an inverse probability censoring weighted c-index are modified to account for discrete risk scores. Simulations are generated to assess the finite sample properties of the concordance probability estimate and the weighted c-index. An application of these measures of discriminatory power to a metastatic prostate cancer risk classification system is examined.

Keywords

C-index; Concordance probability estimate; Discrimination; Inverse probability censoring weight; Risk classification

1. Introduction

Treatment decisions are based on prognosis derived from a set of molecular, pathological, and histological patient factors. A treatment decision algorithm is often determined from a reduced set of important prognostic markers that are categorized to create a risk classification system. The quality of the treatment decision algorithm is tied to the strength of the risk classification.

Discrimination is an important metric in assessing the strength of a risk classification system in survival analysis. When the endpoint is survival time, good discrimination implies that the ordinality of the risk groups predicts the ranking of survival times. If a risk group

Corresponding author: Glenn Heller, ; Email: hellerg@mskcc.org, telephone: 646 735 8112, fax: 646 735 8100

classification can be constructed to confidently delineate long-term survivors from short-term survivors, then it will provide utility in clinical decision making.

The most widely used metric for the global assessment of discrimination is the concordance probability. Its popularity stems from its equivalence to the area under the receiver operating curve when the outcome is binary, and its relation to Kendall's tau and the Goodman-Kruskal gamma for continuous outcomes (Sprent 1989). For survival outcomes, the concordance probability is defined for two hypothetical subjects as $\Pr[R_2 > R_1 | T_1 > T_2]$, where R represents the risk score and T is survival time. The patient-specific risk scores represent the discrete risk groups and may be determined through a regression model via the linear combination $\beta^T X$, with X as a dummy variable. If there are p risk groups, the k^{th} component of the dummy variable X takes the value 1 if the patient is in group k (for k = 1, ..., p-1) and all other components of X are zero. In this paper it is assumed that the proportional hazards model $h(t|X) = h_0(t) \exp(\beta^T X)$ was used to determine patient risk. The application of this model requires that the adequacy of the proportional hazards assumptions have been empirically tested and that the model assumptions cannot be rejected. Two common methods used to test the proportional hazards assumption are developed in Lin et al. (1993) and Grambsch and Therneau (1994).

The outline of this paper is as follows. In Section 2, two previously developed estimates of the concordance probability are presented for continuous risk scores. The estimates, the weighted c-index and the concordance probability estimate (CPE), have derived asymptotic distributions with continuous risk scores. In Section 3, these estimates are adapted for discrete risk scores. The asymptotic distributions of the CPE and the weighted c-index with discrete risk scores are presented. An analytic expression for the asymptotic variance of the CPE is derived. For the weighted c-index, a resampling based estimate of the asymptotic variance is utilized. Simulations are presented in Section 4 to examine the finite sample properties of the estimates. The development of a risk group classification system in prostate cancer is used illustrate the methods in Section 5 and further discussion points are provided in Section 6.

2. Concordance probability with continuous risk scores

It is assumed that that a correctly specified proportional hazards model was used to develop the risk score and that larger values of $\beta^T X$ represent a greater risk of death. The concordance probability, derived from this regression model, is

$$\mathscr{C} = \Pr(\boldsymbol{\beta}^T \boldsymbol{X}_1 > \boldsymbol{\beta}^T \boldsymbol{X}_2 | T_2 > T_1).$$

An early estimate of the concordance probability, adapted for survival analysis, is the cindex (Harrell et al. 1982)

$$\frac{\sum_{i}\sum_{j}I(y_{i}< y_{j})I(\delta_{i}=1)I(\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i}>\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})}{\sum_{i}\sum_{i}I(y_{i}< y_{i})I(\delta_{i}=1)},$$

where y is the minimum of the survival time t and censoring time u, $\delta = I(t - u)$, and $\beta^T X$ is the estimated risk score from the regression model. Although this has been the predominant discrimination statistic used in survival analysis, it has been observed that the statistic is biased and its distribution is a function of the censoring distribution (Gönen and Heller 2005, Uno et al. 2011). Indeed, the limiting value of Harrell's c-index is $\Pr[\beta^T X_1 > \beta^T X_2 | T_1 < T_2, T_1 < U_1 \land U_2]$. Thus, the c-index has the undesirable property that it is influenced by the rate of patient accrual and the length of the study.

To address this limitation, Uno et al. (2011) introduced inverse probability censoring weights to the c-index. The inverse probability censoring weights are incorporated to produce a consistent estimate of the concordance probability when the support of the survival distribution is less than the support of the censoring distribution. The weighted c-index is computed as

$$\frac{\sum_{i}\sum_{j}I(y_{i}< y_{j})I(\delta_{i}=1)I(\hat{\pmb{\beta}}^{T}\pmb{X}_{i}>\hat{\pmb{\beta}}^{T}\pmb{X}_{j})\{\hat{G}(y_{i})\}^{-2}}{\sum_{i}\sum_{j}I(y_{i}< y_{j})I(\delta_{i}=1)\{\hat{G}(y_{i})\}^{-2}},$$

where $\hat{G}(y)$ is the Kaplan-Meier estimated survival function for the censoring time random variable. For this weighted c-index, it is assumed that the censoring random variable is independent of the risk covariates. An extension developed by modeling the censoring distribution as a function of the covariates was proposed by G^{erds} et al. (2013). In general, measures based on inverse probability censoring weights require care as they may be sensitive to late failure times.

If $\{T_i, X_i\}$ are continuous, independent identically distributed random variables, an application of Bayes theorem shows that the concordance probability is also equal to $\Pr[T_2 > T_1 | \boldsymbol{\beta}^T X_1 > \boldsymbol{\beta}^T X_2]$. This form of the concordance probability is used to derive an alternative estimate (Gönen and Heller 2005), based on the proportional hazards specification

$$m(T) = \boldsymbol{\beta}^T \boldsymbol{X} + \varepsilon,$$

where ε is a standard extreme value random variable and m is a monotone function (Kalbfleisch 1978). It follows that

$$\Pr(T_2 > T_1 | \boldsymbol{\beta}^T \boldsymbol{X}_1 > \boldsymbol{\beta}^T \boldsymbol{X}_2) = \frac{\int \int \int [1 + \exp \{\boldsymbol{\beta}^T (\boldsymbol{X}_2 - \boldsymbol{X}_1)\}]^{-1} dF(\boldsymbol{\beta}^T \boldsymbol{X}_1) dF(\boldsymbol{\beta}^T \boldsymbol{X}_2)}{\int \int \int dF(\boldsymbol{\beta}^T \boldsymbol{X}_1) dF(\boldsymbol{\beta}^T \boldsymbol{X}_2)}$$
$$\boldsymbol{\beta}^T \boldsymbol{X}_1 > \boldsymbol{\beta}^T \boldsymbol{X}_2$$

where F is the distribution function of the risk score $\beta^T X$. The concordance probability estimate (CPE) is attained by substituting the partial likelihood estimate for β and the empirical distribution function for F

$$\left. \frac{2}{n(n-1)} \sum_{i < j} \left\{ \frac{I(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_{ji} < 0)}{1 + \exp(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_{ji})} + \frac{I(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_{ij} < 0)}{1 + \exp(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_{ij})} \right\},$$

where $X_{ii} = X_i - X_i$.

3. Concordance probability with discrete risk scores

The concordance probability is adapted for use with discrete risk scores. In general, the creation of risk groups induces a large number of ties in the risk scores. For p groups, with n_k subjects in group k, the proportion of subject pairs in the same group is

$$\frac{\sum_{k=1}^{p} n_k (n_k - 1)}{(\sum_{k=1}^{p} n_k) [(\sum_{k=1}^{p} n_k) - 1]}.$$

Thus for example, a 200 patient study with 50 patients in each of 4 risk groups, nets approximately 25% of the patient pair comparisons from the same group.

3.1 Concordance parameters

There are two paths to modify the concordance probability \mathscr{C} for discrete risk scores. The first is to acknowledge the ambiguity resulting from the inclusion of ties in the risk scores

$$\mathscr{C}_{I} = \Pr(\boldsymbol{\beta}^{T} \boldsymbol{X}_{1} > \boldsymbol{\beta}^{T} \boldsymbol{X}_{2} | T_{2} > T_{1}) + [0.5 \times \Pr(\boldsymbol{\beta}^{T} \boldsymbol{X}_{1} = \boldsymbol{\beta}^{T} \boldsymbol{X}_{2} | T_{2} > T_{1})],$$

and the second is to exclude patient pairs with the same risk score

$$\mathcal{C}_{\scriptscriptstyle E} \! = \! \Pr(\boldsymbol{\beta}^T \boldsymbol{X}_1 \! > \! \boldsymbol{\beta}^T \boldsymbol{X}_2 | T_2 \! > \! T_1, \boldsymbol{\beta}^T \boldsymbol{X}_1 \neq \boldsymbol{\beta}^T \boldsymbol{X}_2).$$

The concordance measures \mathscr{C}_I and \mathscr{C}_E are the survival analogs of Kendall's tau-a and Kendall's tau-b, respectively. Yan and Greene (2008) demonstrate that \mathscr{C}_I is a convex combination of \mathscr{C}_E and 0.5

$$\mathscr{C}_{I} = [\mathscr{C}_{E} \times \{1 - \Pr(\boldsymbol{\beta}^{T} \boldsymbol{X}_{1} = \boldsymbol{\beta}^{T} \boldsymbol{X}_{2})\}] + [0.5 \times \Pr(\boldsymbol{\beta}^{T} \boldsymbol{X}_{1} = \boldsymbol{\beta}^{T} \boldsymbol{X}_{2})].$$

The parameter \mathscr{C}_E is a common concordance measure which ranges from 0.5 to 1.0. In contrast, the inclusion of ties attenuates \mathscr{C}_I toward 0.5, with its upper bound equal to $1-0.5 \times \Pr(\beta^T X_1 = \beta^T X_2)$. Thus, the use of \mathscr{C}_I penalizes the concordance measure, with fewer risk groups incurring a greater penalty.

It is assumed that $\beta^{\widehat{T}}X_1$ $\beta^{\widehat{T}}X_2$ unless $X_1 = X_2$ (patient pairs within the same risk group). Under this assumption, conditional on X, the asymptotic distribution of the estimate of \mathscr{C}_I follows immediately from the asymptotic distribution of the estimated \mathscr{C}_E . As a result, the focus in this paper is on the statistical properties of concordance estimates based on \mathscr{C}_E .

3.2 Concordance parameter estimates

The weighted c-index adapted for discrete risk scores is

$$C_{n,E}(\hat{\boldsymbol{\beta}}) = \frac{\sum_{i} \sum_{j} I(y_i < y_j) I(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_i > \hat{\boldsymbol{\beta}}^T \boldsymbol{X}_j) I(\delta_i = 1) \{\hat{G}(y_i | \hat{\boldsymbol{\beta}}^T \boldsymbol{X}_i) \hat{G}(y_i | \hat{\boldsymbol{\beta}}^T \boldsymbol{X}_j)\}^{-1}}{\sum_{i} \sum_{j} I(y_i < y_j) I(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_i \neq \hat{\boldsymbol{\beta}}^T \boldsymbol{X}_j) I(\delta_i = 1) \{\hat{G}(y_i | \hat{\boldsymbol{\beta}}^T \boldsymbol{X}_i) \hat{G}(y_i | \hat{\boldsymbol{\beta}}^T \boldsymbol{X}_j)\}^{-1}}.$$

This estimate explicitly excludes ties in the risk scores and has been modified to account for possibly unequal follow-up between risk groups, using $\hat{G}(y|\beta^T X)$ as the within group Kaplan-Meier survival function for the censoring time random variable. The weighted cindex is a consistent estimate of \mathscr{C}_E when the right support of the survival distribution is less than the right support of the censoring distribution. If the support of the survival distribution exceeds the support of the censoring distribution, then the estimate is stabilized by truncating the follow-up time at τ so that $G(\tau|\beta^T X)>0$. For the weighted c-index, this is incorporated into the statistic as

$$C_{n,E}(\hat{\boldsymbol{\beta}};\boldsymbol{\tau}) = \frac{\sum_{i}\sum_{j}I(y_{i} < y_{j})I(\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i} > \hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})I(\delta_{i} = 1)I(y_{i} < \tau_{i})\{\hat{\boldsymbol{G}}(y_{i}|\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i})\hat{\boldsymbol{G}}(y_{i}|\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})\}^{-1}}{\sum_{i}\sum_{j}I(y_{i} < y_{j})I(\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i} \neq \hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})I(\delta_{i} = 1)I(y_{i} < \tau_{i})\{\hat{\boldsymbol{G}}(y_{i}|\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i})\hat{\boldsymbol{G}}(y_{i}|\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})\}^{-1}}$$

where τ_i represents the truncation time within the i^{th} subject's risk group.

The asymptotic normality of the weighted c-index for continuous risk scores was demonstrated in Uno et al. (2011). For discrete risk scores, the asymptotic normality is derived in the appendix. This extension follows directly from the work in Cheng et al. (1995) and Fine et al. (1998). Estimation of the asymptotic variance, however, requires density estimation for the within group censoring random variable. To avoid kernel smoothing and bandwidth selection, a stratified by risk group bootstrap standard error estimate of the weighted c-index in the simulations. An alternative resampling approach was used in Uno et al (2011).

The CPE with discrete risk scores is

$$K_{\scriptscriptstyle n,E}(\hat{\boldsymbol{\beta}}) = \frac{\sum_{i} \sum_{j} I(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_i {>} \hat{\boldsymbol{\beta}}^T \boldsymbol{X}_j) \big[1 + \exp(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_{ji})\big]^{-1}}{\sum_{i} \sum_{j} I(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_i {>} \hat{\boldsymbol{\beta}}^T \boldsymbol{X}_j)}.$$

The CPE is a consistent estimate of \mathscr{C}_E when the proportional hazards specification is correct. The consistency and the asymptotic distribution of the CPE, including its asymptotic variance, are derived in the appendix.

Theorem—Under the standard conditions for the proportional hazards model, $n^{1/2}[K_{n,E}(\hat{\beta})] - \mathscr{C}_E$ is asymptotically normal with mean 0.

The CPE is a model based measure of concordance and therefore relies on the correctness of the proportional hazards model. The advantage of using the CPE is that it does not require inverse probability censoring weights, which may be sensitive to late failures. Use of the CPE provides additional insight into the discriminatory power of a risk classification system derived with p risk groups by noting that the CPE is a weighted average of pairwise probabilities, with each element representing the probability the lower risk patient lives longer

$$K_{n,E}(\hat{\pmb{\beta}}) = \frac{\sum_{k=1}^{p} \sum_{l=1}^{p} I[\hat{\beta}_{k} > \hat{\beta}_{l}] n_{k} n_{l} \left[1 + \exp\{\hat{\beta}_{l} - \hat{\beta}_{k}\}\right]^{-1}}{\sum_{k=1}^{p} \sum_{l=1}^{p} I[\hat{\beta}_{k} > \hat{\beta}_{l}] n_{k} n_{l}}.$$

Thus, in addition to computing the overall CPE, we can extract the individual probability $[1+\exp{\{\hat{\beta_I}-\hat{\beta_k}\}}]^{-1}$, that a patient in risk group k lives longer than a patient in risk group l.

4. Simulations

Simulation experiments were conducted to examine the accuracy of the CPE and weighted c-index for a four group risk classification system. The data were generated from Weibull and lognormal regression models. For the Weibull model the survival times were produced from

$$T_i = \exp[-0.5x_{1i} - 0.25x_{2i} - 0.10x_{3i}] \times \varepsilon_i$$

where $X = (x_1, x_2, x_3)^T$ are dummy variables. The errors $\{\varepsilon_i\}$ were independent identically distributed Weibull random variables with scale parameter 1 and shape parameters $\{1.85, 4.1, 7.3, 13.5\}$, which were chosen to produce concordance probabilities equal to $\{0.6, 0.7, 0.8, 0.9\}$. The Weibull regression model satisfies the proportional hazards assumption. For the lognormal model, the log survival times were created using the same covariates and regression parameters. The log of the error random variables $\{\log \varepsilon_i\}$ were independent identically distributed normal random variables with mean 0 and scale parameters $\{0.62, 0.28, 0.16, 0.09\}$, which were again chosen to produce concordance probabilities equal to $\{0.6, 0.7, 0.8, 0.9\}$. The lognormal simulations were constructed to examine the robustness of the CPE and the weighted c-index statistics when the regression coefficients were computed using the proportional hazards model but the data do not satisfy the proportional hazards assumptions.

For each group, independent uniform censoring times $(0, M_k)$, (k = 1, 2, 3, 4), were generated to determine the overall proportion censored. The sample size for each simulation was 200 and the proportion of patients in the four groups was $\{0.1, 0.2, 0.3, 0.4\}$. One thousand simulations were run in each setting to compute the weighted c-index excluding ties $(C_{n,E})$, and the CPE excluding ties $(K_{n,E})$. The results of the simulations are summarized in Tables 1–3.

When proportional hazards is correct, the results shown in Table 1 demonstrate that the CPE is relatively unbiased, except when the concordance probability is small. The bias that

occurs when the regression model is weak is a function of the bias in the Cox regression coefficient estimates in these simulations. For the weighted c-index, in addition to the bias when the CP = 0.60, a bias is present when the censoring proportion is high. The deficiency of the weighted c-index when data are highly censored is likely attributable to the shorter censoring distribution support relative to the survival distribution support; a violation of a key assumption for the weighted c-index stated in Section 2. The weighted c-index with a time truncation, $C_{n,E}(\beta, \tau)$ in Section 3.2, attenuated the weighted c-index and was not helpful in reducing the bias (Table 2). The truncation time was chosen such that within each risk group, $G(\tau|\beta^T X) > 0.10$.

The analytic estimated standard error of the CPE derived in this paper is close to its simulation standard error, and the empirical coverage for the concordance probability based the asymptotic 95% confidence interval $K_{n,E} \pm 1.96 \times \text{se}(K_{n,E})$ is good. As expected, the estimated standard error of the CPE increases as the percent censoring increases and the concordance probability (CP) decreases. For the weighted c-index, the stratified bootstrap resampling standard error and coverage estimate are also accurate.

If the proportional hazards assumption is true, a benefit to using the CPE relative to the weighted c-index is efficiency. The simulation standard errors of $K_{n,E}$ are uniformly smaller than the simulation standard errors of $C_{n,E}$. In terms of mean squared error, the relative efficiency of the CPE to the weighted c-index is always greater than 1 and is as high as 1.4.

For nononproportional hazards data, Table 3 indicates that the CPE is not accurate. The CPE is biased and its coverage may be poor. The inadequacy of the CPE in this case is unsurprising, since its kernel is based on a result stemming from extreme value survival times

$$\Pr[T_2 > T_1 | \boldsymbol{X}_1, \boldsymbol{X}_2] = \frac{1}{1 + \exp\{\boldsymbol{\beta}^T (\boldsymbol{X}_2 - \boldsymbol{X}_1)\}},$$

whereas for normally distributed log survival times

$$\Pr[T_2 > T_1 | \boldsymbol{X}_1, \boldsymbol{X}_2] = 1 - \Phi\left(\frac{\boldsymbol{\beta}^T \boldsymbol{X}_1 - \boldsymbol{\beta}^T \boldsymbol{X}_2}{\sqrt{2}\sigma}\right),$$

with $\Phi(\cdot)$ representing the standard normal distribution function and σ the scale parameter. The results in Table 3 reinforce the use of diagnostics to assess proportionality before applying the CPE to a given set of data. Conversely, the weighted c-index appears to be sufficently robust and provides similar bias and variance accuracy for both error distributions.

5. Prostate cancer example

In recent years, considerable progress in the treatment of metastatic prostate cancer has resulted in new treatments using immunotherapy, hormone therapy, and chemotherapy, with each demonstrating a prolongation of life. As a result, patient prognosis, as determined

through risk groups, can affect the choice of treatment. Importantly, the incorporation of a risk classification system in the treatment decision process requires an understanding of the strength of this system to discriminate clinical outcome. A risk classification system was developed from 148 newly diagnosed metastatic prostate cancer patients. After screening multiple factors, patient risk was based on a categorization of two important biomarkers in this patient population: circulating tumor cells (CTC) and lactate dehydrogenase (LDH) (Scher et al. 2009). CTC is a blood-based assay that provides information on the accumulation of tumor cells in the peripheral blood. LDH is a marker of cell turnover and is considered an indirect marker of tumor burden. The finding that five or more CTCs at baseline is associated with shorter survival times has been found in multiple metastatic solid tumor populations (Danila et al. 2011). The upper limit of normal for LDH was defined as 250 units per liter as determined by the central laboratory used for this study. As a result of these previously determined normal/abnormal ranges, the risk classification system derived from the two biomarkers was:

Low risk		CTC < 5	
Intermediate risk	CTC	5 LDH	250
High risk	CTC	5 LDH	> 250

The Kaplan-Meier estimates of survival were well separated by risk and are depicted in Figure 1. The median survival times in the three groups were: 23.6 months (low risk), 13.6 months (intermediate risk), and 8.2 months (high risk). A proportional hazards model was generated using this risk classification and a test of the proportional hazards assumption (Grambsch and Therneau 1994) provided insufficient evidence that the proportional hazards assumption was violated (p= 0.293). From the results of the proportional hazards model, the estimated log relative risk of death was 0.860 (se=0.304) and 1.584 (se = 0.280) for the intermediate and high risk groups respectively, compared to the low risk group. This risk classification is clearly prognostic, but p-values alone, which were both <0.005, are insufficient to assess the strength of this classification. In general, p-values are sensitive to the sample size of the data, with a large number of failures producing a downward influence on the p-value. In order to ascertain the strength of this classification system in discriminating patient risk, the concordance probability estimate and weighted c-index were computed.

The CPE and weighted c-index excluding ties were computed to assess the discrimination strength of this risk group classification. The CPE and weighted c-index were equal to 0.753 (se=0.036) and 0.761 (se = 0.060), respectively. These results indicate that the three group staging system provides good discriminatory power and may be used to assist in the treatment decision process.

When patient pairs in the same risk group were included in the discrimination measures, the CPE was equal to 0.659 and the weighted c-index was equal to 0.666, indicating the the derived metastatic prostate cancer risk groups were not sufficiently strong at discriminating

survival rates. However, the overall proportion of pairwise comparisons with patients in the same risk group was 0.30, which explains the attenuation of these discrimination measures.

The CPE and weighted c-index provide a measure of the overall discriminatory power of the risk group classification system. In addition to this overall measure of strength, the CPE can also be used to estimate the discriminatory power between individual risk groups. Using the proportional hazards specification

$$\Pr[T_2 > T_1 | \boldsymbol{X}_1, \boldsymbol{X}_2] = [1 + \exp\{\boldsymbol{\beta}^T (\boldsymbol{X}_2 - \boldsymbol{X}_1)\}]^{-1},$$

and substituting $\hat{\beta}$, the probability is 0.83 that a patient in the low risk group will survive longer than a patient in the high risk group. The probability is 0.70 that a patient in the low risk group survives longer than a patient in the intermediate risk group, and approximately the same probability holds for a patient in the intermediate risk group relative to a patient in the high risk group. This illustrates that although the overall CPE is 0.753, the discriminatory strength between individual groups varies. As noted in Section 3, the overall CPE is a weighted average of these individual measures.

The discriminatory power of a risk classification system should be validated before the system is put into practice. An independent cohort of 97 metastatic prostate cancer patients (Danila et al. 2014) was used to validate the estimates of the concordance probability. The CPE and weighted c-index with ties excluded was 0.772 and 0.783, respectively. These results are comparable to the development study results and confirm the good discriminatory power of the CTC+LDH risk classification system in this patient population.

6. Discussion

The interpretation of the concordance probability differs depending on whether ties in the risk groups are included or excluded. The inclusion of patient pairs within the same risk group reduces the concordance probability (\mathscr{C}_{p}) parameter toward the null value 0.50 and is therefore affected by the distribution of the subjects across risk groups. In contrast, the parameter (\mathscr{C}_{E}), based on the exclusion of patient pairs within the same risk group is unaffected by the distribution of patients across groups.

The most commonly applied measure of discrimination is the unweighted c-index. An advantage of the c-index is that it is nonparametric and may be useful if the proportional hazards assumptions do not hold. We previously showed that with continuous risk scores the unweighted c-index had a bias that was a function of the percent censored (Gönen and Heller 2005). We have adapted the weighted c-index (Uno et al. 2011) to the case when patient pairs in the same risk group are excluded in the calculation. Our simulations indicate that the inverse probability censoring weighted c-index excluding ties is robust to the violation of proportional hazards.

Under proportional hazards, the semiparametric CPE excluding ties was more efficient than the nonparametric weighted c-index. Both statistics were relatively unbiased, but at high censoring rates the weighted c-index demonstrated some bias. In general, time truncation is

used to reduce bias for an inverse probability censoring weight statistic, but our simulations demonstrated a lack of improvement uisng the time truncated weighted c-index. In contrast, the CPE is a model based measure of discrimination. It does not require inverse probability censoring weights and therefore does not require a stabilizing truncation point. It does, however, assume that the proportional hazards relation is correct.

Clinical risk classification systems are ubiquitous in medicine and their usefulness in clinical decision making is a function of their discriminatory power. Risk classification systems differ in their discrimination capability and a measure to assess the strength of a classification system should be applied. Our findings indicate that if the proportional hazards model is employed to develop the classification system, and diagnostics are used to confirm the proportional assumptions, then the CPE is an efficient measure of discrimination. If an alternative approach to risk classification is employed, then the nonparametric weighted c-index produces an accurate assessment of discrimination.

R code to compute the CPE and its standard error, with the option to include or exclude ties in the risk scores is available in the R package CPE. For the weighted c-index, the R package pec enables the user the option to include or exclude ties in the risk scores.

Appendix

The asymptotic distribution for the CPE excluding ties

It is assumed that the proportional hazards model $h(t|\mathbf{X}) = h_0(t) \exp(\boldsymbol{\beta}_0^T \mathbf{X})$ is the correct specification for the data and that the standard conditions for the asymptotic normality of $n^{1/2}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_0)$ apply (Andersen and Gill, 1982).

Consistency of the CPE excluding ties: $K_{n,E}(\hat{\beta}) \stackrel{p}{\to} \mathscr{C}_E$ where

$$K_{n,E}(\hat{\boldsymbol{\beta}}) = \frac{\sum_{i} \sum_{j} I(\hat{\boldsymbol{\beta}}^{T} \boldsymbol{X}_{i} > \hat{\boldsymbol{\beta}}^{T} \boldsymbol{X}_{j}) [1 + \exp(\hat{\boldsymbol{\beta}}^{T} \boldsymbol{X}_{ji})]^{-1}}{\sum_{i} \sum_{j} I(\hat{\boldsymbol{\beta}}^{T} \boldsymbol{X}_{i} > \hat{\boldsymbol{\beta}}^{T} \boldsymbol{X}_{j})}$$

and
$$\mathcal{C}_E = \Pr[\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 > \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2 | T_2 > T_1, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 \neq \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2]$$

Using the proportional hazards specification provided in Section 2 and the law of large numbers,

$$K_{n,E}(\hat{\boldsymbol{\beta}}) \xrightarrow{p} \Pr[T_2 > T_1 | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 > \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2].$$

Note that the concordance probability excluding ties may be written as

$$\mathscr{C}_{E} = \Pr[T_{2} > T_{1} | \boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{1} > \boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{2}] \times \frac{\Pr[\boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{1} > \boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{2}]}{\Pr[T_{2} > T_{1}, \boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{1} \neq \boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{2}]}. \quad (A.1)$$

Now using the independence between subjects, the following relations hold

$$\Pr[\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 > \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2] = \frac{\Pr[\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 \neq \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2]}{2} \quad (A.2)$$

$$\frac{1}{2} = \Pr[T_2 > T_1 | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 \neq \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2] \Pr[\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 \neq \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2] + \Pr[T_2 > T_1 | \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 = \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2] \Pr[\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 = \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2]$$
(A. 3)

Substituting (A.2) into (A.3), it follows that

$$\Pr[T_2 > T_1, \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 \neq \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2] = \Pr[\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_1 > \boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{X}_2] \quad (A.4)$$

and substitution of (A.4) into the denominator of (A.1) proves the consistency of the CPE.

Asymptotic distribution of $n^{1/2}$ [$K_{n,E}(\hat{\beta}) - \mathcal{C}_E$]

To prove the asymptotic normality of the CPE, it is first shown that if

$$|\boldsymbol{\beta}_0^T \boldsymbol{X}_{ij}| > \varepsilon > 0$$
, (A.5)

then $n^{1/2}K_{n,E}(\beta)$ is asymptotically equal to

$$n^{1/2} \tilde{K}_{n,E}(\hat{\boldsymbol{\beta}}) = \frac{n^{-3/2} \sum_{i} \sum_{j} I(\boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{ij} > 0) [1 + \exp(\hat{\boldsymbol{\beta}}^{T} \boldsymbol{X}_{ji})]^{-1}}{n^{-2} \sum_{i} \sum_{j} I(\boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{ij} > 0)}$$

where within the indicator functions the estimate $\hat{\beta}$ is replaced with the true regression coefficient β_0 and $X_{ij} = X_i - X_j$. Note that in the discrete covariate case there are a finite number of covariate values X_{ij} . As a result, (A.5) is not a strong assumption.

The asymptotic equality is demonstrated by considering the cases $X_{ij} = 0$ and $X_{ij} = 0$ separately.

If
$$X_{ij} = 0$$
, then clearly $I(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_{ij} > 0) = I(\boldsymbol{\beta}_0^T \boldsymbol{X}_{ij} > 0)$.

If X_{ij} 0, then using the consistency of $\hat{\beta}$ and assumption (A.5), for *n* sufficiently large,

$$\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_{ij} > v > 0$$
 and $I(\hat{\boldsymbol{\beta}}^T \boldsymbol{X}_{ij} > 0) = I(\boldsymbol{\beta}_0^T \boldsymbol{X}_{ij} > 0)$ for all \boldsymbol{X}_{ij} . Therefore, for n large,

$$n^{1/2}K_{n,E}(\hat{\boldsymbol{\beta}})=n^{1/2}\tilde{K}_{n,E}(\hat{\boldsymbol{\beta}}).$$

A Taylor expansion for the asymptotically equivalent CPE produces

$$n^{1/2}\left[\tilde{K}_{n,E}(\hat{\boldsymbol{\beta}})-\mathscr{C}_{E}\right]=n^{1/2}\left[\tilde{K}_{n,E}(\boldsymbol{\beta}_{0})-\mathscr{C}_{E}\right]+\left(\frac{\partial\tilde{K}_{n,E}}{\partial\boldsymbol{\beta}}\right)^{\mathrm{T}}\left[n^{1/2}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_{0})\right]+o_{p}(1).$$

The partial derivative $K_{n,E}$ β is asymptotically constant. Since $n^{1/2}(\hat{\beta} - \beta_0)$ has asymptotic mean zero conditional on X, it is asymptotically independent of $n^{1/2}[K_{n,E}(\beta_0) - \mathscr{C}_E]$. Therefore the asymptotic variance of $K_{n,E}(\hat{\beta})$ is

$$\operatorname{Var} \ \left\{ n^{1/2} \left[\tilde{K}_{n,E}(\pmb{\beta}_0) - \mathscr{C}_E \right] \right\} + \left(\frac{\partial \tilde{K}_{n,E}}{\partial \pmb{\beta}} \right)^{\operatorname{T}} \operatorname{Var}[\hat{\pmb{\beta}}] \left(\frac{\partial \tilde{K}_{n,E}}{\partial \pmb{\beta}} \right).$$

The individual components of the asymptotic variance can be estimated as follows. In each case, the substitution of $\hat{\beta}$ for $\hat{\beta}_0$ provides a consistent estimate.

The $Var(\hat{\beta})$ is estimated from the inverse of the second derivative of the partial likelihood.

The partial derivative evaluated at β_0 is equal to

$$\frac{\partial \tilde{K}_{n,E}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -\frac{\sum_{i} \sum_{j} I(\boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{ij} > 0) \boldsymbol{X}_{ji} \exp(\boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{ji}) [1 + \exp(\boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{ji})]^{-2}}{\sum_{i} \sum_{j} I(\boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{ij} > 0)}.$$

The first term, $n^{1/2} [K_{n,E}(\beta_0) - \mathscr{C}_E]$, may be approximated by the U-statistic

$$\pi^{-1}n^{-3/2} \sum_{i} \sum_{j} I(\boldsymbol{\beta}_{0}^{T}\boldsymbol{X}_{ij} > 0) \left\{ \left[1 + \exp(\boldsymbol{\beta}_{0}^{T}\boldsymbol{X}_{ji}) \right]^{-1} - \mathscr{C}_{\scriptscriptstyle{E}} \right\}$$

where
$$\pi = \lim_{n \to \infty} n^{-2} \sum_{i} \sum_{j} I(\boldsymbol{\beta}_{0}^{T} \boldsymbol{X}_{ij} > 0)$$

The asymptotic variance of this U-statistic is

$$\pi^{-2}n^{-3}\sum_{i}\sum_{j}\sum_{k\neq j}[u_{ij}+u_{ji}][u_{ik}+u_{ki}].$$

where

$$u_{ij} = I(\boldsymbol{\beta}_0^T \boldsymbol{X}_{ij} > 0) \left\{ \left[1 + \exp(\boldsymbol{\beta}_0^T \boldsymbol{X}_{ji}) \right]^{-1} - \mathscr{C}_E \right\}.$$

Combining these results provides the estimated asymptotic variance of the CPE excluding ties.

Asymptotic distribution of $n^{1/2} [C_{n,E}(\beta) - \mathcal{C}_E]$

From Section 3.2, this expression may be written as

$$\frac{n^{-3/2}\sum_{i}\sum_{j}\delta_{i}I(y_{i}< y_{j})I(\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i}\neq\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})\left\{\hat{G}(y_{i}|\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i})\hat{G}(y_{i}|\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})\right\}^{-1}\left[I(\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i}>\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})-\mathscr{C}_{E}\right]}{n^{-2}\sum_{i}\sum_{j}\delta_{i}I(y_{i}< y_{j})I(\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i}\neq\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})\left\{\hat{G}(y_{i}|\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i})\hat{G}(y_{i}|\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})\right\}^{-1}}\\ =\psi^{-1}n^{-3/2}\sum_{i}\sum_{j}\frac{\delta_{i}I(y_{i}< y_{j})I(\boldsymbol{\beta}_{0}^{T}\boldsymbol{X}_{i}\neq\boldsymbol{\beta}_{0}^{T}\boldsymbol{X}_{j})[I(\boldsymbol{\beta}_{0}^{T}\boldsymbol{X}_{i}>\boldsymbol{\beta}_{0}^{T}\boldsymbol{X}_{j})-\mathscr{C}_{E}]}{G(y_{i}|(\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{i})G(y_{i}|(\hat{\boldsymbol{\beta}}^{T}\boldsymbol{X}_{j})})}+o_{p}(1)$$

where $\psi = \lim_{n \to \infty} n^{-2}$ $_{i}$ $_{j} \delta_{i} I(y_{i} < y_{j}) I(\beta^{\widehat{T}} X_{i} \quad \beta^{\widehat{T}} X_{j}) \hat{G}^{-1}(y_{i} \beta^{\widehat{T}} X_{i}) \hat{G}^{-1}(y_{i} \beta^{\widehat{T}} X_{j}).$

Letting

$$e_{ij} = \delta_i I(y_i < y_i) I(\boldsymbol{\beta}_0^T \boldsymbol{X}_i \neq \boldsymbol{\beta}_0^T \boldsymbol{X}_i) [I(\boldsymbol{\beta}_0^T \boldsymbol{X}_i > \boldsymbol{\beta}_0^T \boldsymbol{X}_i) - \mathscr{C}_E]$$

and Taylor expanding $\hat{\beta}$ around $\hat{\beta}_0$.

$$\begin{split} n^{1/2} \left[C_{n,E}(\hat{\boldsymbol{\beta}}) - \mathcal{C}_E \right] &= \psi^{-1} n^{-3/2} \sum_i \sum_j \frac{e_{ij}}{G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_i) G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_j)} + \\ \psi^{-1} n^{-3/2} \sum_i \sum_j \frac{e_{ij}}{G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_i) G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_j)} \left[\frac{G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_i) - \hat{G}(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_i)}{G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_i)} + \frac{G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_j) - \hat{G}(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_j)}{G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_j)} \right] + \\ \psi^{-1} n^{-3/2} \sum_i \sum_j \left\{ \left[\frac{\partial}{\partial \boldsymbol{\beta}} \frac{e_{ij}}{G(y_i | \boldsymbol{\beta}^T \boldsymbol{X}_i) G(y_i | \boldsymbol{\beta}^T \boldsymbol{X}_j)} \right]^T \right\} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) + o_p(1) \end{split}$$

The second term may be rewritten using the martingale representation theorem as

$$2n^{-1/2} {\sum_{k=1}^n} {\int_{t=0}^\infty} \frac{q_{{\pmb X}_k}(t)}{\theta_{{\pmb X}_k}(t)} dM_{{\pmb X}_k}(t)$$

where

$$q_{\boldsymbol{X}_k}(t) = \lim_{n \to \infty} \sum_{i} \sum_{j} \left\{ \frac{e_{ij}}{G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_i) G(y_i | \boldsymbol{\beta}_0^T \boldsymbol{X}_j)} \right\} \left\{ I(y_j \geq t) [I(\boldsymbol{X}_i = \boldsymbol{X}_k) + I(\boldsymbol{X}_j = \boldsymbol{X}_k)] \right\}$$

$$\theta_{\boldsymbol{X}_k}(t) = \lim_{n \to \infty} \sum_{j} I(y_j \ge t, \boldsymbol{X}_j = \boldsymbol{X}_k)$$

$$M_{X_k}(t) = I(y_k \le t, \delta_k = 0) - \int_{u=0}^{t} I(y_k \ge u) d\Lambda_{X_k}(u)$$

and Λ_{X_k} is the cumulative hazard of the censoring random variable belonging to group X_k (Cheng et al. 1995). Fine et al. 1998).

It follows that $n^{1/2} [C_{n,E}(\beta) - \mathcal{C}_E]$

$$\begin{split} =& \psi^{-1} n^{-3/2} \sum_{i} \sum_{j} \frac{e_{ij}}{G(y_{i}|\boldsymbol{\beta}_{0}^{T}\boldsymbol{X}_{i})G(y_{i}|\boldsymbol{\beta}_{0}^{T}\boldsymbol{X}_{j})} + 2\psi^{-1} n^{-1/2} \sum_{k=1}^{n} \int_{t=0}^{\infty} \frac{q_{\boldsymbol{X}_{k}}(t)}{\theta_{\boldsymbol{X}_{k}}(t)} dM_{\boldsymbol{X}_{k}}(t) \\ +& \psi^{-1} \left\{ n^{-2} \sum_{i} \sum_{j} \left[\frac{\partial}{\partial \boldsymbol{\beta}} \frac{e_{ij}}{G(y_{i}|\boldsymbol{\beta}^{T}\boldsymbol{X}_{i})G(y_{i}|\boldsymbol{\beta}^{T}\boldsymbol{X}_{j})} \right]^{T} \right\} n^{1/2} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_{0}) + o_{p}(1) \end{split}$$

Each component is asymptotically normal with mean zero. Analytic estimation of the asymptotic variance, however, requires density estimation of the censoring random variable. Alternatively, a stratified bootstrap resampling approach is utilized to estimate the asymptotic variance.

References

- Andersen PK, Gill RD. Cox's regression model for counting procsses: A large sample study. Annals of Statistics. 1982; 10:1100–1120.
- Cheng SC, Wei LJ, Ying Z. Analysis of transformation models with censored data. Biometrika. 1995; 82:835–845.
- Cox DR. Regression models and life tables (with Discussion). Journal of the Royal Statistical Society, Series B. 1972; 34:187–220.
- Cox DR. Partial likelihood. Biometrika. 1975; 62:269-276.
- Danila DC, Fleisher M, Scher HI. Circulating tumor cells as biomarkers in prostate cancer. Clinical Cancer Research. 2011; 17:3903–3912. [PubMed: 21680546]
- Danila DC, Anand A, Schultz N, Heller G, Wan M, Sung CC, Dai C, Khanin R, Fleisher M, Lilja H, Scher HI. Analytic and clinical validation of a prostate cancer-enhanced messenger RNA detection assay in whole blood as a prognostic biomarker for survival. European Urology. 2014; 65:1191–1197. [PubMed: 23954088]
- Fine JP, Ying Z, Wei LJ. On the linear transformation model for censored data. Biometrika. 1998; 85:980–986.
- Gerds TA. Prediction Error Curves for risk prediction models in survival analysis. R package version 2.4.2. 2014
- Gerds TA, Kattan MW, Schumacher M, Yu C. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. Statistics in Medicine. 2013; 32:2173–2184. [PubMed: 23172755]
- Gönen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. Biometrika. 2005; 92:965–970.
- Grambsch PM, Therneau TM. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. Biometrika. 1994; 81:515–526.
- Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. Journal of the American Medical Association. 1982; 247:2543–2546. [PubMed: 7069920]
- Kalbfleisch JD. Likelihood methods and nonparametric tests. Journal of the American Statistical Association. 1978; 73:167–170.
- Lin DY, Wei LJ, Ying Z. Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals. Biometrika. 1993; 80:557–572.
- Mo Q, Gönen M, Heller G. CPE: Concordance Probability Estimates in Survival Analysis. R package version 1.4.3. 2012

R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011. URL http://www.R-project.org/

- Scher HI, de Bono JS, Fleisher M, Pienta KJ, Raghavan D, Heller G. Circulating tumour cells as prognostic markers in progressive, castration-resistant prostate cancer: a reanalysis of IMMC38 trial data. Lancet Oncology. 2009; 10:233–239. [PubMed: 19213602]
- Sprent, P. Applied Nonparametric Statistical Methods, Second edition. London: Chapman and Hall; 1989.
- Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Statistics in Medicine. 2011; 30:1105–1117. [PubMed: 21484848]
- Yan G, Greene T. Investigating the effects of ties on measures of concordance. Statistics in Medicine. 2008; 27:4190–4206. [PubMed: 18384187]

Heller and Mo

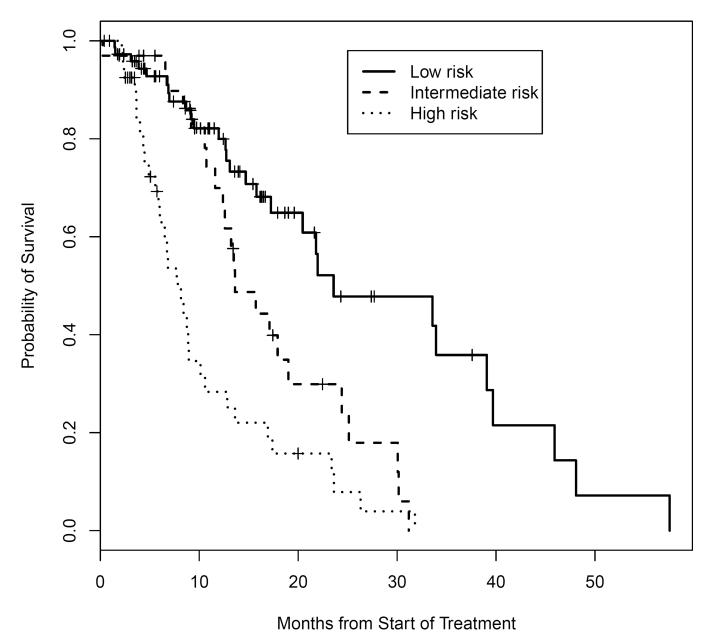


Figure 1.Kaplan-Meier estimates of survival based on clinical risk.

Author Manuscript

Author Manuscript

Proportional hazards data

RE	1.212	1.151	1.030	1.204	1.385	1.066	1.059	1.350	1.233	1.093	1.164	1.418	1.331	1.305	1.231	1.366
$\operatorname{cov}(K_{n,E})$	0.958	0.956	0.958	0.941	0.958	0.963	0.955	0.955	0.963	0.953	0.947	0.946	0.949	0.957	0.940	0.928
$\operatorname{cov}(C_{n,E})$	0.953	0.938	0.927	0.929	0.944	0.955	0.935	0.929	0.937	0.949	0.944	0.950	0.928	0.939	0.936	0.942
$\operatorname{se}(K_{n,E})$ avg	0.027	0.032	0.041	0.058	0.024	0.028	0.035	0.051	0.020	0.023	0.028	0.041	0.014	0.015	0.019	0.026
$\operatorname{se}(C_{n,E})$ avg	0.028	0.030	0.037	0.056	0.025	0.029	0.035	0.052	0.021	0.023	0.030	0.045	0.015	0.016	0.020	0.031
$K_{n,E}$ sim se	0.025	0.029	0.036	0.046	0.022	0.026	0.034	0.046	0.019	0.022	0.027	0.040	0.013	0.014	0.018	0.027
$C_{n,E}$ sim se	0.028	0.031	0.037	0.054	0.026	0.027	0.035	0.053	0.021	0.023	0.029	0.045	0.015	0.016	0.020	0.031
$K_{n,E}$ avg	0.605	0.605	0.609	0.624	0.702	0.703	0.702	0.701	0.800	0.800	0.801	0.797	0.900	0.901	0.901	0.899
$C_{n,E}$ avg	0.602	0.606	0.607	0.618	0.700	0.701	0.698	0.693	0.798	0.800	0.797	0.784	0.900	0.899	0.900	0.894
Prop cen	0.00	0.25	0.50	0.75	0.00	0.24	0.49	0.75	0.50	0.24	0.50	0.75	0.00	0.24	0.50	0.75
ච	9.0	9.0	9.0	9:0	0.7	0.7	0.7	0.7	8.0	8.0	8.0	0.8	6:0	6.0	6.0	6.0

CP = Concordance Probability; Prop cen = proportion censored $C_{n,E}$ = Weighted c-index excluding ties; $K_{n,E}$ = CPE excluding ties

 $avg=average; sim\ se=simulation\ standard\ error; cov=coverage$ $RE=Mean\ square\ error\ efficiency\ of\ the\ CPE\ relative\ to\ the\ weighted\ c-index$

Heller and Mo

Table 2

Weighted c-index with time truncation; proportional hazards data

$\mathrm{cov}(C_{n,E}(\tau))$	0.953	0.938	0.930	0.945	0.944	0.955	0.933	0.914	0.937	0.949	0.944	0.935	0.928	0.939	0.936	0.945
$\operatorname{cov}(C_{n,E})$	0.953	0.938	0.927	0.929	0.944	0.955	0.935	0.929	0.937	0.949	0.944	0.950	0.928	0.939	0.936	0.942
$\operatorname{se}(C_{n,E}(\tau))$ avg	0.028	0.030	0.037	0.055	0.025	0.029	0.035	0.055	0.021	0.023	0.030	0.050	0.015	0.016	0.020	0.033
$\left \begin{array}{c} \operatorname{se}(C_{n,E}) \\ \operatorname{avg} \end{array}\right $	0.028	0.030	0.037	0.056	0.025	0.029	0.035	0.052	0.021	0.023	0.030	0.045	0.015	0.016	0.020	0.031
$C_{n,E}(\tau)$ sim se	0.028	0.031	0.038	0.053	0.026	0.027	0.036	0.056	0.021	0.023	0.030	0.051	0.015	0.016	0.020	0.033
$C_{n,E}$ sim se	0.028	0.031	0.037	0.054	0.026	0.027	0.035	0.053	0.021	0.023	0.029	0.045	0.015	0.016	0.020	0.031
$C_{n,E}(au)$	0.602	909.0	0.604	0.612	0.700	0.701	969.0	0.677	0.798	0.800	0.796	0.769	0.900	0.899	0.900	0.888
$C_{n,E}$ avg	0.602	909.0	0.607	0.618	0.700	0.701	0.698	0.693	0.798	0.800	0.797	0.784	0.900	0.899	0.900	0.894
Prop cen	0.00	0.25	0.50	0.75	0.00	0.24	0.49	0.74	0.00	0.24	0.50	0.75	0.00	0.24	0.49	0.75
CP C	9.0	9.0	9.0	0.6	0.7	0.7	0.7	0.7	0.8	8.0	8.0	0.8	6.0	6.0	6.0	0.9

 $CP = Concordance\ Probability;\ Prop\ cen = proportion\ censored$

 $C_{n,E}$ = Weighted c-index excluding ties

 $C_{n,E}(au)= {
m Time-truncated weighted } c{
m -index excluding ties, } G(y|m{\beta}^T X)>0.10$

avg = average; sim se = simulation standard error; cov = coverage

Page 18

Table 3

Log normal survival time data

$\operatorname{cov}(C_{n,E})$ $\operatorname{cov}(K_{n,E})$	0.953 0.940	0.934 0.975	0.920 0.950	0.903 0.873	0.931 0.836	0.934 0.916	0.946 0.942	0.911 0.906	0.950 0.691	0.948 0.829	0.929 0.936	0.921 0.926	0.939 0.583	0.944 0.725	0.933 0.858	
$\begin{vmatrix} \sec(K_{n,E}) \\ avg \end{vmatrix}$	0.027	0.032	0.040	0.057	0.024	0.028	0.035	0.049	0.021	0.023	0.028	0.038	0.015	0.017	0.019	
$\operatorname{se}(C_{n,E})$ avg	0.030	0.031	0.037	0.055	0.027	0.031	0.037	0.052	0.021	0.025	0.031	0.046	0.014	0.017	0.021	
$K_{n,E}$ sim se	0.026	0.028	0.034	0.048	0.026	0:030	0.035	0.045	0.023	0.026	0:030	0.040	0.018	0.020	0.024	
$\begin{vmatrix} C_{n,E} \\ \text{sim se} \end{vmatrix}$	0.028	0.032	0.037	0.055	0.028	0.031	0.038	0.056	0.020	0.024	0.032	0.050	0.015	0.017	0.020	
$K_{n,E}$ avg	0.590	0.601	0.614	0.641	0.675	0.686	0.700	0.726	0.769	0.779	0.789	0.807	0.873	0.877	0.885	
$\begin{vmatrix} C_{n,E} \\ ext{avg} \end{vmatrix}$	0.602	0.605	0.612	0.629	0.699	0.701	0.700	0.705	0.800	0.801	0.800	0.795	0.900	0.899	0.900	
Prop cen	0.00	0.24	0.50	0.75	0.00	0.25	0.50	0.75	0.00	0.24	0.49	0.75	0.00	0.24	0.49	1
ච	9.0	9.0	9.0	0.6	0.7	0.7	0.7	0.7	8.0	8.0	8.0	0.8	6.0	6.0	6.0	

CP = Concordance Probability; Prop cen = proportion censored $C_{n,E}$ = Weighted c-index excluding ties; $K_{n,E}$ = CPE excluding ties avg = average; sim se = simulation standard error; cov = coverage