Concordance for Survival Time Data: Fixed and Time-Dependent Covariates and Possible Ties in Predictor and Time

Walter K. Kremers

Technical Report Series #80
April 2007

Department of Health Sciences Research
and The William J. von Liebig Transplant Center
Mayo Clinic
Rochester, Minnesota

This work was supported in part by a grant from the National Institutes of Health (DK-34238).

Concordance for Survival Time Data: Fixed and Time-Dependent Covariates and Possible Ties in Predictor and Time

Walter K Kremers
Division of Biostatistics, Department of Health Sciences Research
and The William J. von Liebig Transplant Center
Mayo Clinic and Mayo Clinic College of Medicine
200 First Street SW
Rochester, MN 55905

April 17, 2007

Support: This work was supported in part by a grant from the National Institutes of Health (DK-34238).

Abstract

Concordance, or synonymously the C-statistic, is a valuable measure of model discrimination in analyses involving survival time data. Generally it is defined based upon baseline covariates and with only limited consideration of ties. Here we provide a definition of concordance in the case of survival data; we allow for time-dependent covariates with the counting process data representation and account for ties in the covariates and times. This definition reduces to standard definitions of concordance in the case of fixed covariates and in the absence of ties. This definition provides a general measure with similar intuition as earlier measures. It also allows not only for comparison between models with time-dependent covariates but also for comparison of discrimination between models with and without time-dependent covariates. Asymptotic standard errors and methods for obtaining approximate confidence intervals are also provided.

Key Words

Discrimination; Concordance; C-Statistic; C-Index; Survival; Time-dependent covariate; Variance estimate; Cox proportional hazards

1. Introduction

A common and useful measure of model discrimination is concordance or synonymously the C-statistic, which is applicable to any ordered outcome. In general, consider selecting random pairs of patients, and for each pair note whether the model correctly predicts order, e.g., a higher model score for the better result. Concordance is then the fraction of pairs for which the model is correct. A completely random prediction would have a concordance of 0.5, a perfect rule a concordance of 1.

Concordance is most familiar from logistic regression, where it is also known as the area under the receiver operating curve. It can also be described for survival data while allowing for censoring and no distributional assumption need be made for motivation or calcualtion[1, 2]. Concordance for survival data can be described based upon all data pairs obtained by selecting two individuals from the sample. Each element of each pair has a survival or censor time, an indicator of event or censoring and a predictor (risk) score. If, based upon the information contained in a pair, one can determine which individual first incurred an event, then we say this pair is evaluable. Depending on censoring, some pairs may not be evaluable, in particular, if one individual is censored before the other individual incurs an event, which includes the case where both individuals are censored. Concordance then is the fraction of all evaluable pairs where the predictor score correctly predicts (is greater) for the individual with the earlier event. If there were no censoring, this would be essentially Kendall's Tau. This measure is closely related to concordance from the logistic model[3, 4] and both are widely used in the medical literature to measure model discrimination[5-7]. Generally concordance in survival analysis is described for the case with fixed covariates known at baseline though it has also been described for the case of time dependent covarites[8, 9]. Here, we describe concordance for the case with time dependent covariates and also allow for ties in the predictor and ties in event times. Note, for calculations, covariates need only be known at the time of each event. A variance formula recently described for concordance[10] is restricted to applications where there are no ties in the predictor or event times. Here, we

derive a variance formula that not only allows for time-dependent covariates but also allows for ties in times and predictors, thus allowing for derivation of confidence intervals for a broad class of applications. We also derive an asymptotic variance sometimes used in practice for which the derivation is not well documented.

2. Definition of Concordance for time-dependent covariates

For the definition of concordance for survival time data with time-dependent covariates, consider a random sample of independent and identically distributed individuals with survival data, where for each individual we observe the earlier of an event time or a censor time and that the event and censor times are independent, and where we have knowledge of a covariate which is a function of time, and that the covariate function can be determined at each event time when an individual is at risk. Though not required for calculation of concordance, we also consider this covariate function to be left continuous to assure the covariate is predictive and not simply correlative.

From the sample, index those individuals that incurred events by i, and denote the corresponding event time by t_i. We will then define a concordance measure based upon counts derived from those individuals who incurred events. In particular, considering all individuals except i at risk at time t_i and who do not have an event at time t_i let C_i be the count of individuals with predictor score at time t_i less than the predictor score for individual i at time t_i, let D_i be the count of individuals with predictor score at time t_i greater than the predictor score for individual i at time t_i and let P_i be the count of individuals with predictor score at time t_i equal to the predictor score for individual i at time t_i. Further, again considering all individuals except i at risk at time t_i, let T_i be the count of individuals who too have an event at time t_i. Then, C_i, D_i, P_i and T_i may be regarded as the count of pairs formed with individual i at time t_i concordant, discordant, tied in prediction and tied in time. Note, from this construct

individuals who are tied in both predictor and time are counted in T_i and not in P_i. Conceptually we can consider concordance for i at time t_i as [C_i + ½ P_i]/[C_i + D_i + P_i] following Harrell[1] or as [C_i + ½ P_i + ¼ T_i]/[C_i + D_i + P_i + ½ T_i] following Therneau in his formulation of the survConcordance function in the S-Plus® software (S-Plus® 7, Insightful Corp, Seattle, WA, 2005). Summing over all patients with events we define the concordance for the whole sample as

1.1) $$\left(\sum_i C_i + \tfrac{1}{2}\sum_i P_i\right) \Big/ \left(\sum_i C_i + \sum_i D_i + \sum_i P_i\right)$$

following Harrell or as

1.2) $$\left(\sum_i C_i + \tfrac{1}{2}\sum_i P_i + \tfrac{1}{4}\sum_i T_i\right) \Big/ \left(\sum_i C_i + \sum_i D_i + \sum_i P_i + \tfrac{1}{2}\sum_i T_i\right)$$

following Therneau. The formulation 1.2) is also suggested by considering Kendall's Tau where the contribution is intermediate if individuals are tied on either (though not both) of two variables[2]. According to the counting mechanism described above individuals tied in time are counted twice, once when considering individual i as the reference and again when considering individual j as the reference, and therefore the T_i terms are multiplied by a half when contrasted with the P_i terms.

In this description of concordance there is explicitly no reference to counting process data representation. Still, the comparison of every individual at risk each time there is an event requires knowledge of covariates at each event time when an individual is at risk and this is provided by the counting process data representation[11]. Note, as is generally required of time dependent covariates, the covariates must be "known" at the event times, that is they must be a function of the current or past and not the future. For actual calculations only values of covariates for individuals in the risk set when events occur enter into the calculations. In practice it will sometimes be convenient to record in the data representation just the covariates of individual at risk each time there was an event rather than recording all changes in covariates.

3. Consistency with Earlier Definitions of Concordance using baseline or time-independent covariates

To show that the above definition of concordance is consistent with earlier definitions consider the above definition for the case when a) the time-dependent covariates are degenerately constant, that is identical to the baseline predictors and b) all individuals are continually at risk from time 0 to either their event or censor time. The earlier definitions of concordance consider the set of all possible pairs possibly drawn from the sample, the counts of pairs concordant, discordant, tied in prediction and tied in time. Definition 1.1 and 1.2 are also based upon counts of pairs possibly drawn from the sample. The only difference is that 1.1 and 1.2 essentially index by event times rather than by observation number. Because the index only serves to assure that all pairs are properly counted, the counts are invariant to the index and definitions 1.1 and 1.2 are the same as those of Harrell and Therneau. The minor difference between the two counting mechanisms concerns ties in time, where the index based upon observation number does not lead to double counting, but where our counting mechanism based in part upon event times double counts. Adjusting for this as we have, the 1.1 and 1.2 are consistent with earlier definitions of concordance.

4. An Alternate Formulation, Variance and Confidence Intervals

In defining concordance for time-dependent data we essentially took the event times as an index. Consideration of the formulae though shows that an alternate expression can be given, which is more similar to the usual description of concordance for the case without time-dependent covariates. Concordance can be expressed as a function of all pairs of individuals where each pair contributes 1, ½ or 0 to the numerator if the pair is concordant, tied or discordant. Here concordance for the pair is inferred if it can be determined which one of the pair first had an event and if at the time of that first event the time-dependent predictor of the individual with an event was greater than the individual's time-dependent predictor with the later event. Ties and discordance are defined

analogously. Letting C_{ij}, D_{ij}, P_{ij} and S_{ij} be the contributions based upon pair (i, j), 1.1) and 1.2) can be based upon sums of C_{ij}, D_{ij}, P_{ij} and S_{ij}, in particular

1.3) $(\sum_{i<j} C_{ij} + ½ \sum_{i<j} P_{ij})/(\sum_{i<j} C_{ij} + \sum_{i<j} D_{ij} + \sum_{i<j} P_{ij})$

1.4) $(\sum_{i<j} C_{ij} + ½ \sum_{i<j} P_{ij} + ½ \sum_{i<j} T_{ij})/(\sum_{i<j} C_{ij} + \sum_{i<j} D_{ij} + \sum_{i<j} P_{ij} + \sum_{i<j} T_{ij})$

represent 1.1) and 1.2). From these formulae it is a straight forward matter, though tedious, to obtain variances and covariances for and between the terms $\sum_{i<j} C_{ij}$, $\sum_{i<j} D_{ij}$, $\sum_{i<j} P_{ij}$ and $\sum_{i<j} T_{ij}$ taking special care of the dependence between terms with a single common subscript. For the description of Cov($\sum_{i<j} X_{ij}$, $\sum_{i<j} Y_{ij}$) where X_{ij} and Y_{ij} are generic for any of C_{ij}, D_{ij}, P_{ij} and S_{ij}, observe that

$$\mathrm{E}[\sum_{i<j} X_{ij} * \sum_{i<j} Y_{ij}] = \mathrm{E}[\sum_{i<j} X_{ij}Y_{ij}] + \mathrm{E}[\sum_{\substack{((i<j)\&(k<l))\&\\((i\neq k)|(j\neq l))\&\\((i=k)|(i=l)|\\(j=k)|(j=l))}} X_{ij}Y_{kl}] + \mathrm{E}[\sum_{\substack{((i<j)\&(k<l))\&\\((i\neq k)\&(i\neq l)\&\\(j\neq k)\&(j\neq l))}} X_{ij}Y_{kl}]$$

The first expectation to the right of the equals sign consists of the product of terms X_{ij} and Y_{kl} sharing both subscript values; the second expectation consists of the product terms where exactly one of the subscript values is shared; the third term consists of the product of terms where none of the subscript values is shared. By independence of X_{12} and Y_{34} the above expression is the same as

$$Q\,\mathrm{E}[X_{12}Y_{12}] + R\,\mathrm{E}[X_{12}Y_{13}] + S\,\mathrm{E}[X_{12}]\mathrm{E}[Y_{12}]$$

where Q, R and S are the number of terms in the respective summations. Because of the identity $\sum_{i=1}^{I} i = I(I+1)/2$, $Q = (n-1)n/2$, where n denotes the sample size. For derivation of R observe that X_{ij} can be paired with $(n-i-1)$ terms Y_{kl} where $k=i$ and $l\neq j$, and $(i-1)$ Y_{kl} terms where $l=i$ and $k\neq l$, that is there are $(n-2)$ Y_{kl} terms sharing the subscript i and not j. Considering the index j shows an additional $(n-2)$ Y_{kl} terms sharing one subscript in common with X_{ij}

for a total of $2(n-2)$ terms. As there are $(n-1)n/2$ X_{ij} terms, $R = (n-2)(n-1)n$. As there are a total of Q^2 product terms, $S = (Q^2 - Q - R)$, and

1.5) $\text{Cov}(\sum_{i<j} X_{ij}, \sum_{i<j} Y_{ij}) = Q\,\text{E}[X_{12}Y_{12}] + R\,\text{E}[X_{12}Y_{13}] - (Q+R)\,\text{E}[X_{12}]\,\text{E}[Y_{12}]$

Recognizing that for any *i<j*, that at most one of the C_{ij}, D_{ij}, P_{ij} and S_{ij} can take the value 1, and since all terms can only take the value 0 or 1, we have that

$$\text{Var}(\sum_{i<j} X_{ij}) = Q\,\text{E}[X_{12}] + R\,\text{E}[X_{12}X_{13}] - (Q+R)\,\text{E}[X_{12}]^2$$

and for X different from Y

$$\text{Cov}(\sum_{i<j} X_{ij}, \sum_{i<j} Y_{ij}) = R\,\text{E}[X_{12}Y_{13}] - (Q+R)\,\text{E}[X_{12}]\text{E}[Y_{12}]$$

Substituting the $\text{E}[X_{12}]$ and $\text{E}[X_{12}Y_{13}]$ by the sample moments yields estimates for $\text{Var}(\sum_{i<j} X_{ij})$ and $\text{Cov}(\sum_{i<j} X_{ij}, \sum_{i<j} Y_{ij})$. Computations of the moments are dominated by $\text{E}[X_{12}Y_{13}]$ for which there are $R = (n-2)(n-1)n$ terms involved. Thus these computations are of order n^3 and may be extensive for large samples.

By linearity of the numerator and denominator in 1.4)

$$\text{Var}(\sum_{i<j} C_{ij} + \tfrac{1}{2}\sum_{i<j} P_{ij} + \tfrac{1}{2}\sum_{i<j} T_{ij}) = \text{Var}(\sum_{i<j} C_{ij}) + \tfrac{1}{4}\text{Var}(\sum_{i<j} P_{ij}) + \tfrac{1}{4}\text{Var}(\sum_{i<j} T_{ij})$$
$$+ \text{Cov}(\sum_{i<j} C_{ij}, \sum_{i<j} P_{ij}) + \text{Cov}(\sum_{i<j} C_{ij}, \sum_{i<j} T_{ij}) + \tfrac{1}{2}\text{Cov}(\sum_{i<j} P_{ij}, \sum_{i<j} T_{ij})$$

$$\text{Var}(\sum_{i<j} C_{ij} + \sum_{i<j} D_{ij} + \sum_{i<j} P_{ij} + \sum_{i<j} T_{ij}) =$$
$$\text{Var}(\sum_{i<j} C_{ij}) + \text{Var}(\sum_{i<j} D_{ij}) + \text{Var}(\sum_{i<j} P_{ij}) + \text{Var}(\sum_{i<j} T_{ij})$$
$$+ 2\,\text{Cov}(\sum_{i<j} C_{ij}, \sum_{i<j} D_{ij}) + 2\,\text{Cov}(\sum_{i<j} C_{ij}, \sum_{i<j} P_{ij}) + 2\,\text{Cov}(\sum_{i<j} C_{ij}, \sum_{i<j} T_{ij})$$
$$+ 2\,\text{Cov}(\sum_{i<j} D_{ij}, \sum_{i<j} P_{ij}) + 2\,\text{Cov}(\sum_{i<j} D_{ij}, \sum_{i<j} T_{ij}) + 2\,\text{Cov}(\sum_{i<j} P_{ij}, \sum_{i<j} T_{ij})$$

$$\text{Cov}(\sum_{i<j} C_{ij} + \tfrac{1}{2}\sum_{i<j} P_{ij} + \tfrac{1}{2}\sum_{i<j} T_{ij}\ ,\ \sum_{i<j} C_{ij} + \sum_{i<j} D_{ij} + \sum_{i<j} P_{ij} + \sum_{i<j} T_{ij}) =$$
$$\text{Var}(\sum_{i<j} C_{ij}) + \tfrac{1}{2}\text{Var}(\sum_{i<j} P_{ij}) + \tfrac{1}{2}\text{Var}(\sum_{i<j} T_{ij})$$

$$+ \text{Cov}(\sum_{i<j} C_{ij}, \sum_{i<j} D_{ij}) + 1\ ½\ \text{Cov}(\sum_{i<j} C_{ij}, \sum_{i<j} P_{ij}) + 1\ ½\ \text{Cov}(\sum_{i<j} C_{ij}, \sum_{i<j} T_{ij})$$

$$+ ½\ \text{Cov}(\sum_{i<j} D_{ij}, \sum_{i<j} P_{ij}) + ½\ \text{Cov}(\sum_{i<j} D_{ij}, \sum_{i<j} T_{ij}) + \text{Cov}(\sum_{i<j} P_{ij}, \sum_{i<j} T_{ij})$$

Based upon these variances and covariances one may construct approximate confidence intervals for concordance using Fieller's theorem[12] which is based upon the identity for the bivariate normal (X,Y) of

1.6) $\text{P}((X - rY)^2 > z_{\alpha/2}^2(\text{Var}(X) - 2r\,\text{Cov}(X,Y) + r^2\,\text{Var}(Y)) = \alpha$

for $r = \mu_X / \mu_Y$ and where $z_{\alpha/2}$ is the α/2 quantile of the standard normal distribution. As the numerator and denominator in 1.4) are approximately normal, making the appropriate substitutions for X, Y, $\text{Var}(X)$, $\text{Var}(Y)$ and $\text{Cov}(X,Y)$ in 1.6) and solving the inequality for r yields confidence intervals for concordance with level very nearly $100(1-\alpha)\%$.

More simply, based upon Taylor series expansion

$\text{Var}(X/Y) \approx (\mu_X/\mu_Y)^2\ (\text{Var}(X)/\mu_X^2 + \text{Var}(Y)/\mu_Y^2 - 2*\text{Cov}(X,Y)/(\mu_X\mu_Y))$

This approximation can be used to obtain an approximate standard deviation for 1.4), and in turn an approximate $100(1-\alpha)\%$ confidence interval with limits 1.4) $\pm z_{\alpha/2}$ times the standard deviation. Similar derivations may be made for 1.3).

Computationally simpler than the "direct" methods described above for derivation of variances, standard deviations and confidence intervals are resampling methods like the bootstrap and jackknife. The bootstrap method will likely provide more accurate confidence intervals for the true concordance. Confidence intervals based on the jackknife estimate of standard error, however, will generally be computationally less intense. Applicability of the jackknife is supported by the recognition that 1.3) and 1.4) are functions of U-statistics[13]. Though the resampling methods may be simple to program their run times may be long compared to the direct methods.

Computationally simpler than the jackknife itself is an approximate jackknife, based upon an approximation of $\left(\widehat{\theta_{(i)}} - \widehat{\theta}\right)$. Let α and β be the numerator and denominator terms in the description of concordances 1.3 or 1.4

but with the summation over all $i \neq j$ instead of $i < j$. Note, α and β involve double counting. Let α_{ij} be the contribution to the numerator obtained when comparing individuals i and j and similarly β_{ij} for the denominator. Let $\alpha_{i\bullet} = \sum_{j \neq i} \alpha_{ij}$ (only j varies in the summation) and $\beta_{i\bullet} = \sum_{j \neq i} \beta_{ij}$. Then, since $\alpha = \sum_i \alpha_{i\bullet}$ and $\beta = \sum_i \beta_{i\bullet}$, we can express concordance as $\hat{\theta} = \alpha/\beta$. If $\widehat{\theta_{(i)}}$ is concordance calculated leaving out individual i then, by the double counting in $\hat{\theta}$,

$$\widehat{\theta_{(i)}} = \frac{(\alpha - 2\alpha_{i\bullet})}{(\beta - 2\beta_{i\bullet})}$$

indicating the jackknife variances may be obtained in practice without a complete rederivation for each subsample. Further, by Taylor's series expansion,

$$\left(\widehat{\theta_{(i)}} - \hat{\theta}\right) \approx -(1/\beta)(2\alpha_{i\bullet} - \alpha) + (\alpha/\beta^2)(2\beta_{i\bullet} - \beta)$$

and

$$\sum_i \left(\widehat{\theta_{(i)}} - \hat{\theta}\right)^2 \approx$$

$$4\sum_i \left[(1/\beta^2)(\alpha_{i\bullet} - \alpha)^2 + (\alpha^2/\beta^4)(\beta_{i\bullet} - \beta)^2 - 2(\alpha/\beta^3)(\alpha_{i\bullet} - \alpha_{\bullet\bullet})(\beta_{i\bullet} - \beta_{\bullet\bullet})\right]$$

$$= 4(\alpha/\beta)^2 \sum_i \left[(\alpha_{i\bullet} - \alpha)^2/\alpha^2 + (\beta_{i\bullet} - \beta)^2/\beta^2 - 2(\alpha_{i\bullet} - \alpha)(\beta_{i\bullet} - \beta)/(\alpha\beta)\right]$$

$$= 4(\alpha/\beta)^2 \sum_i \left[\alpha_{i\bullet}{}^2/\alpha^2 + \beta_{i\bullet}{}^2/\beta^2 - 2\alpha_{i\bullet}\beta_{i\bullet}/(\alpha\beta)\right]$$

yielding the approximate jackknife estimate of variance. Considering nonparametric correlations, Quade[14] describes an asymptotic variance of the same general form. In this article, though, Quade cites an earlier work for justification of the variance formulae[15] which I have not been able to obtain. This general variance formula is also used by Frank Harrell for calculation of the concordance variance in his R/S-Plus rcorr.cens functions Ratfor/Fortran service routine[16]. The approximate jackknife, at least as programmed by myself, requires minimally less computation and programming than the jackknife, though this difference is very small.

This jackknife asymptotic variance can also be motivated considering the general properties of U-statistics directly as described by Hoeffding[17] and Lehmann[18]. Lehmann also describes an identity (formula A.202, page 368) relating Cov(X_{ij} , X_{ik}) and $Var(X_{ij} \mid data\,for\,indivdual\,i)$ which can be used to describe variance formula using a direct method with calculations of the order n^2, though we do not consider this further here. Since all $\alpha_{i\bullet}$ and $\beta_{i\bullet}$ can be easily obtained numerically using a double loop over i and j, the approximate jackknife provides a simple and intuitive method for variance estimation.

5. An application

In an earlier work[19], we studied the ability of the Model for End-Stage Liver Disease (MELD) score, which is a function of blood creatinine level, bilirubin level and prothrombin time, to discriminate survival in patients with advanced liver disease. Because of the low number of censors and specific interest in the time points studied of 1 week, 3 months and 1 year, we used the logistic model instead of modeling actual survival times and found concordance measures in the range of 73% to 95%. In a subsequent work[20] we considered the improvements in estimating survival obtained by including the updated MELD score based upon updated measures of creatinine, bilirubin and prothrombin time in patients, using a Cox proportional hazards regression model with time-dependent covariates. Although a major finding of the paper was that patients who had had multiple labs drawn were at an increased risk of mortality (likely a selection effect), the time-dependent (updated) MELD score was also a stronger predictor of survival than baseline MELD score. Using follow up information on 861 patients including 80 deaths, baseline MELD score was associated with patient survival with a concordance of 79.18% (using 1.2). Using Fieller's theorem the approximate 95% confidence interval is (73.82%, 84.26%). The estimated standard error was 2.59% using the direct method, 2.60% by the approximate jackknife and 2.64% by the jackknife. Note, we would not advocate interpretation of the point or confidence estimates to the accuracy of 0.01% but

provide this level of precision here for the comparison of the different methods for estimating confidence intervals.

Using the time-dependent or most current MELD score information to predict survival, 3861 records were required to describe the changes in MELD scores during follow up, and concordance was 81.95% (using 1.2) for an increase of about 2.8%. Using Fieller's theorem the approximate 95% confidence interval is (77.03%, 86.70%). For the time-dependent MELD score, the estimated standard error of concordance was 2.40% by the direct method, 2.41% by the approximate jackknife and 2.44% by the jackknife.

For both of these examples we see that the approximate jackknife was nearer the direct estimate than the jackknife. The standard error estimates were relatively insensitive to which of the three methods was used in estimation. Still, all of the methods based upon an estimated standard error of concordance gave confidence intervals ever so slightly shorter than those derived using Fieller's theorem. Examples can also be constructed where the methods based upon the estimated standard error of concordance give longer confidence intervals than those derived using Fieller's theorem.

For the calculation of standard errors and confidence intervals, Fieller's theorem and the direct method took 4 minutes and the jackknife about 15 seconds when run on the same platform. An attempt was made to program efficiently but not necessarily optimally. The coding for the approximate jackknife and for the jackknife was less involved than that of the direct method. Programs were written in SAS® (SAS Institute Inc., Cary, NC, 2003) and are available from the author.

6. Extensions

The general definition of concordance is motivated by the need to account for time dependent covariates. In creating the definition, though, we have essentially employed the counting process representation of the survival data[11]. Thus our definition of concordance can also be applied to data where individuals are entering and leaving the risk set, irrespective of whether individuals'

covariates are time dependent or time independent. This is especially important when considering alternate time scales in the Cox model such as patient age, or calendar time, which may be indicated when these alternate time scales are stronger determinates of survival than the more typical construct time since study entry[21].

The general definition of concordance can be easily adapted to the case of competing risks survival analysis. Here, we could base a single overall measure of concordance upon the number of concordant and discordant pairs considering each event of any type comparing the predictor of that event type with the predictors of other individuals at risk at the time of the respective event. Being a reduction in dimensions, this overall measure of concordance could not provide the same information as the collection of the individual concordances but still could serve as a single summary measure of model discrimination.

7. Discussion

When modeling survival outcomes using time-dependent covariates, a natural model is the Cox proportional hazards model[22, 23]. For the calculation of concordance, however, no model is required. Calculation of concordance depends solely on the survival information and the predictor scores at the event times. In practice though, there are often multiple predictors of survival and they are fit using a Cox model. The Cox model then yields estimates of $\mathrm{X}\beta$, which serves as a summary score for each patient's relative hazard, and the estimate can be used for the calculation of concordance. When fitting a Cox model with time-dependent covariates, statistical softwares commonly provide estimates of $\mathrm{X}\beta$ making the calculation of concordance a straight forward matter for most anticipated applications.

The general definition of concordance can be re-expressed in terms of ranks, which may offer an additional insight to the behavior of concordance for survival data. First, considering the case when there are no ties for individual i and time t_i, the rank of the score for individual i, amongst the set at risk at time

t_i, minus 1, equals the number of concordant pairs. That is, if R_i denotes the rank of the individual i with event at t_i, then $C_i = (R_i - 1)$, Further if N_i is the size of the risk set at time t_i, then $C_i + D_i$, the number of pairs for comparison with individual i at time t_i, equals $(N_i - 1)$ and concordance can be represented as

$$\sum_i (R_i - 1) / \sum_i (N_i - 1)$$

Making appropriate definitional modifications to account for ties, we see that concordance, the probability of a model correctly predicting which of two randomly chosen individuals will first have an event, is closely related to the average rank of individuals with events relative to the average risk set size.

8. Conclusion

With the general definition of concordance for survival data considered here, one is able to significantly broaden the types of survival data for the description of concordance. We have a means for measuring model discrimination which is itself directly interpretable and also allows direct comparison between all models to which the definition applies. In particular, the measure can be used to compare between two models with time-dependent covariates or between two models where one model is based upon baseline covariates and the other is based upon baseline and time-dependent covariates and allowing for ties. This general definition of concordance, together with its variance estimates, should be an asset when evaluating discrimination in survival models.

9. Acknowledgement

I thank Frank Harrell for bringing earlier works of Dana Quade to my attention and Terry Therneau for sharing with me so many of his insights on survival analysis.

References

1. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* Feb 28 1996;15(4):361-387.

2. Schemper M. Analysis of associations with censored data by generalized Mantel and Breslow tests and generalized Kendall Correlation. *Biometrical Journal.* 1984;26:309-318.

3. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* Apr 1982;143(1):29-36.

4. McNeil BJ, Hanley JA. Statistical approaches to the analysis of receiver operating characteristic (ROC) curves. *Med Decis Making.* 1984;4(2):137-150.

5. Stephenson AJ, Scardino PT, Eastham JA, et al. Postoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. *J Clin Oncol.* Oct 1 2005;23(28):7005-7012.

6. Kim HL, Seligson D, Liu X, et al. Using tumor markers to predict the survival of patients with metastatic renal cell carcinoma. *J Urol.* May 2005;173(5):1496-1501.

7. Ripoll C, Banares R, Rincon D, et al. Influence of hepatic venous pressure gradient on the prediction of survival of patients with cirrhosis in the MELD Era. *Hepatology.* Oct 2005;42(4):793-801.

8. Antolini L, Boracchi P, Biganzoli E. A time-dependent discrimination index for survival data. *Stat Med.* Dec 30 2005;24(24):3927-3944.

9. Schemper M, Stare J. Explained variation in survival analysis. *Stat Med.* Oct 15 1996;15(19):1999-2012.

10. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med.* Jul 15 2004;23(13):2109-2123.

11. Therneau TM, Grambsch PM. *Modeling survival data: extending the Cox model*. New York: Springer-Verlag; 2000.

12. Finney DJ. *Probit Analysis, Third Edition*. London: Cambridge University Press; 1971.

13. Miller RG. The Jackknife - a review. *Biometrika.* 1974;61:1-17.

14. Quade D. Nonparametric partial correlation. In: Blalock HMJ, ed. *Measurement in the Social Sciences: Theories and Strategies*. Chicago: Aldine; 1974:369-398.

15. Quade D. *Nonparametric Partial Correlation.* Amsterdam: Mathematisch Centrum; 1971.

16. Frank Harrell, personal communication.

17. Hoeffding W. A Class of statistics with asymptotically normal distribution. *he Annals of Mathematical Statistics.* 1948;19:293-325.

18. Lehmann WL. *Nonparametrics*. San Francisco: Holden-Day, Inc.; 1975.

19. Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology.* Feb 2001;33(2):464-470.

20. Bambha K, Kim WR, Kremers WK, et al. Predicting survival among patients listed for liver transplantation: an assessment of serial MELD measurements. *Am J Transplant.* Nov 2004;4(11):1798-1804.

21. Clayton D, Hills M. *Statistical Models in Epidemiology. Section 30.4, p. 302*. New York: Oxford University Press Inc.; 1993.

22. Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society Series B (Methodological).* 1972;34:187-220.

23. Kalbfleisch J, Prentice R. *The Statistical Analysis of Failure Time Data*. New York: Wiley; 2002.