

Cumulative/dynamic ROC curve estimation under interval censorship

Susana Díaz-Coto^{1*}, Pablo Martínez-Camblor²
Norberto Octavio Corral-Blanco¹

¹Department of Statistics, University of Oviedo, Oviedo, Asturias, Spain

²Geisel School of Medicine, Dartmouth College, Hanover, NH, USA

Abstract

The Receiver Operating Characteristic (ROC) curve is a graphical tool commonly used to assess the discriminatory ability of continuous markers in binary classification problems. Different extensions of the ROC curve have been proposed in the prognosis context, where the characteristics in study are time-dependent events. Perhaps the most direct generalization is the so-called cumulative/dynamic (C/D) ROC curve. The main particularity when dealing with the C/D ROC curve estimation is the presence of incomplete information. Several approximation methods addressing this censoring problem have been suggested in the statistical literature, most of them focused on the right-censored case. Interval censorship arises naturally from those studies where subjects undergo periodical follow-ups. They may miss a scheduled appointment and the exact event times are only known to fall in a certain range. A new approach for estimating the C/D ROC curve under the particular scheme of interval censorship is presented in this work. Its finite-sample behavior is studied via Monte Carlo simulations on two different scenarios. Results suggest that the proposed approximation is better than the existing one in terms of absolute error. Its direct application is illustrated in the real-world data set which motivated this research. The uniform strong consistency and a suitable R function for its practical implementation are provided as appendices.

Keywords: (Bio)marker, C/D ROC curve, Interval censorship, Mixed subjects approach, Time-dependent outcome.

*Correspondence to: Susana Díaz-Coto, susanacoto@gmail.com

1 Introduction

Biomedical researchers frequently deal with problems regarding diagnosis and prognosis. Both involve the tasks of classifying subjects and making decisions. The former aims to identify the presence of certain condition according to the symptoms observed. The latter refers to the prediction of a future event (frequently the appearance of a condition of interest) based on symptoms measured at baseline. The Receiver Operating Characteristic (ROC) curve [Zhou et al., 2002, Pepe, 2003] is a graphical tool routinely used to study the discriminatory ability of continuous markers, on binary outcomes. It plots the sensitivity, Se , or proportion of individuals with the condition under study who are correctly classified as positive, against the complement of the specificity, $1 - Sp$, or proportion of subjects wrongly declared positive. Assuming (without loss of generality) that larger values of the marker are associated with higher probabilities of being positive, the ROC curve depicts the points $\{1 - Sp(c), Se(c)\}$ for each $c \in \mathbb{R}$. Equivalently, it is defined by

$$\mathcal{R}(p) = Se(Sp^{-1}(1 - p)) = 1 - F(G^{-1}(1 - p)), \quad p \in [0, 1],$$

where F and G are the cumulative distribution functions (CDF) of the marker variable referred to the positive and negative subjects, respectively; $Sp^{-1}(\cdot) = \inf\{x : Sp(x) \geq \cdot\}$ and analogously $G^{-1}(\cdot) = \inf\{x : G(x) \geq \cdot\}$. The classification groups defining a subject as positive are subsets in the way (c, ∞) with $c \in \mathbb{R}$ ($c = G^{-1}(1 - p)$). Besides, the area under the ROC curve, $AUC (= \int \mathcal{R}(p)dp)$, is the most commonly used index of global diagnostic accuracy [Fluss et al., 2005]. It has several appealing properties which give it advantage on its competitors. Martínez-Cambor and Pardo-Fernández [2019] proved that the area under the standard ROC curve can be read as the probability that, for two randomly selected subjects, one negative and one positive, there exists an eligible classification subset such that both subjects are correctly classified. Under particular conditions, this interpretation can be extended to more general ROC curves, for instance, to the gROC curve proposed by Martínez-Cambor et al. [2017] (in this case, the classification subsets are in the form $(-\infty, a] \cup [b, \infty)$ with $a \leq b \in \mathbb{R}$).

Prognosis is closely related to time-to-event outcomes. The interest lies in determining when the event in study is likely to appear, in order to identify groups of individuals at risk. Different ROC curve extensions have been proposed to accommodate these time-

dependent outcomes (see, for instance, Etzioni et al. [1999] or Pepe et al. [2008]). Among them, the so-called cumulative/dynamic (C/D) ROC curve [Heagerty et al., 2000], based on cumulative sensitivity and dynamic specificity (defined in next section), is perhaps the most direct ROC curve generalization. Moreover, prognosis study designs involve a follow-up period and the presence of incomplete information or censorship is frequent. The most common mechanism that can lead to censored data is right censoring: at the end of the follow-up, some of the enrolled subjects may not have experienced the event of interest or this follow-up may be missed for others for reasons unrelated to the study. In any case, the actual status for those individuals remains unknown. Several C/D ROC curve estimation methods have also been proposed in the specialized literature addressing the right censoring problem.

Different circumstances give rise to another censoring scheme called interval censorship. This is the case, for example, of clinical trials, where individuals are not constantly monitored. They are scheduled periodically to evaluate their condition and some of the appointments may be missed. The exact event time is not observed and is only known to lie between two consecutive revisions. C/D ROC curve estimation under this particular scheme is scarce in the statistical literature. It should be noticed that, once set the time point of interest, t , subjects with unknown status are those whose observable interval contains the point t . There exist a naive approximation [Li and Ma, 2011] where these undefined individuals are removed from the original sample and two model-based approaches for another frameworks (semi competing risks), provided by Jacqumin-Gadda et al. [2016].

In this paper we propose a new C/D ROC curve estimator under interval censorship, the so-called *Mixed subjects* approach. Undefined individuals are considered as both, positive and negative, with certain probabilities. The cumulative sensitivity and dynamic specificity are accordingly approximated in terms of these probabilities.

Rest of the paper is organized as follows. In Section 2 the problem of generalizing the ROC curve to incorporate the time-dimension is revisited. Different types of time-dependent sensitivity and specificity are introduced and defined here. In Section 3 we briefly review the C/D ROC curve estimation problem under the usual right censoring context. The interval censored scenario is presented and a revision of the estimation proposals for this particular scenario is also provided. In Section 4, we present the *Mixed subjects* approach for the estimation of the C/D ROC curve under interval censorship. A non-exhaustive revision

of the available R packages (free at CRAN; www.r-project.com) for computing different estimation methods of C/D ROC curves is exposed in Section 5. Monte Carlo simulations computed for studying the performance of our proposal are described in Section 6 and an illustrative application to a real-world dataset is shown in Section 7. In Section 8 we expose our conclusions. Full details about the strong uniform consistency of the proposed estimator and an R function for its practical implementation are provided as appendices.

2 Time-dependent ROC curves

Prognosis processes refer to a prediction of a future event of interest (outcome), frequently based on the observed symptoms measured at baseline. In this context, a subject will be positive or negative depending on when s/he is evaluated. Different definitions emerge for positive and negative individuals giving rise to proper extensions of sensitivity and specificity and the corresponding time-dependent ROC curves. Among those extensions we focus on the C/D ROC curve, based on the so-called cumulative sensitivity, Se_t^C and dynamic specificity, Sp_t^D , introduced by Heagerty et al. [2000]. For a fixed point of time, t , the former gives the proportion of positive subjects correctly identified as such (considering as positive all individuals suffering the studied event before or at time t). The latter gives the proportion of negative subjects correctly classified, considering as negative those subjects who may experience the event in study beyond t . Formally, once set the time t ,

$$Se_t^C(c) = \mathcal{P}\{X > c \mid T \leq t\},$$

$$Sp_t^D(c) = \mathcal{P}\{X \leq c \mid T > t\},$$

where T and X stand for the time-to-event and marker variables, respectively and $c \in \mathbb{R}$ is the so-called threshold or cutt-off point of the marker. The C/D ROC curve is given by

$$\mathcal{R}_t^{C/D}(p) = Se_t^C([1 - Sp_t^D]^{-1}(p)), \quad p \in [0, 1],$$

where, for the fixed $t \in \mathbb{R}^+$, $[1 - Sp_t^D]^{-1}(\cdot) = \inf\{x : 1 - Sp_t^D(x) \leq \cdot\}$. At time t , the area under the C/D ROC curve has the expression

$$AUC^{C/D}(t) = \int_0^1 \mathcal{R}_t^{C/D}(p) dp.$$

The AUC can be interpreted as the probability that given two randomly chosen subjects, one suffering the event of interest prior to t and the other one beyond t , the marker value will be greater for the first than for the second one.

Other time-dependent ROC curves are the Incident/Dynamic (I/D) ROC curve [Heagerty and Zheng, 2005] and the Incident/Static (I/S) ROC curve [Etzioni et al., 1999, Slate and Turnbull, 2000]. At a time point t , the I/D ROC curve displays the incident sensitivity against the complement of the dynamic specificity. The incident sensitivity gives the proportion of subjects declared positive by the marker among those experiencing the event in study exactly at the considered time t .

On the other hand, at time t , the I/S ROC curve plots the pairs formed by the incident sensitivity and the complement of the static specificity. The later stands for the proportion of subjects classified as negative by the marker among those event-free during a fixed follow-up period $(0, t^*)$. The time t^* is considered long enough to observe the event so the negative subjects are also known as ‘long term survivors’.

3 C/D ROC curve estimation

3.1 Right censorship

Longitudinal studies are frequently subject to censorship. In practice, the right censoring is the most common pattern. The only information provided by a right-censored individual ensures that the event time is greater than the observation time. Mathematically, it is assumed that for the i th individual ($1 \leq i \leq n$) there is a time-to-event, t_i , a censoring time, c_i and it is only observed $z_i = \min\{t_i, c_i\}$ (we have then the status indicator $\delta_i = 1$ if $z_i = t_i$; $\delta_i = 0$ if $z_i = c_i$). Sometimes independence between time-to-event and censoring variables is also assumed (independent right censoring). Figure 1 shows an example of the status of five individuals considering two times, t_1 and t_2 , throughout the follow-up period. At t_1 , all subjects are negative except number 3, who has already experienced the event in study before that time, so it is positive. On the other hand, at t_2 , subjects 2 and 5 are still negative while individual 1 has become positive. No affirmations can be made regarding the number 4, whose follow-up has been missed. It is a right-censored observation. Besides, it should be noticed that, at the end of the follow-up, subject 5 neither has experienced

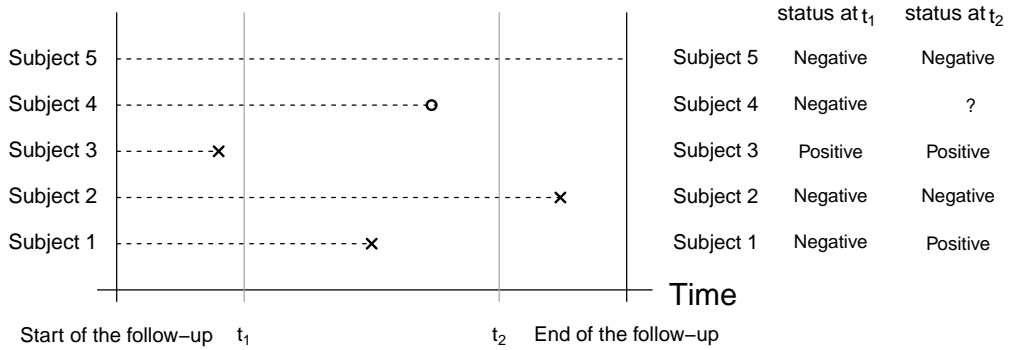


Figure 1: Examples of the status of five subjects at times t_1 and t_2 . A cross represents the event time and a circle means that the follow-up has been missed for the observation.

the target event nor has dropped out from the study, however its status remains unknown. It is another right-censored observation.

Different methods have been proposed for the estimation of the C/D ROC curve under right censorship. An illustrative revision can be found in Blanche et al. [2013]. The authors also provided a brief discussion of the applicability of the methods according to whether or not censorship depends on the marker. Once set a time of interest t , the simplest proposal is the *naive* method, where observations censored before t are removed from the original sample. The C/D ROC curve is approximated by the observed true-positive and false-negative proportions, over the remaining individuals. Obviously, a loss of information takes place and the estimator may lead to biases. There exist other procedures accounting to the information provided for the censored observations before t . Among them, those provided by Heagerty et al. [2000], Chambles and Diao [2006], Uno et al. [2007], Hung and Chiang [2010], Blanche et al. [2013], Wolf et al. [2011] and the smooth estimation methodology suggested by Martínez-Cambor and Pardo-Fernández [2018].

Finally, we highlight an interesting C/D ROC curve estimation method proposed by Martínez-Cambor et al. [2016]. Censored observations before the considered time t are treated as mixed subjects: they are partially allocated to both, positive and negative groups. We will refer this method as the *Mixed subjects* approach and it will be exposed in detail in the next section.

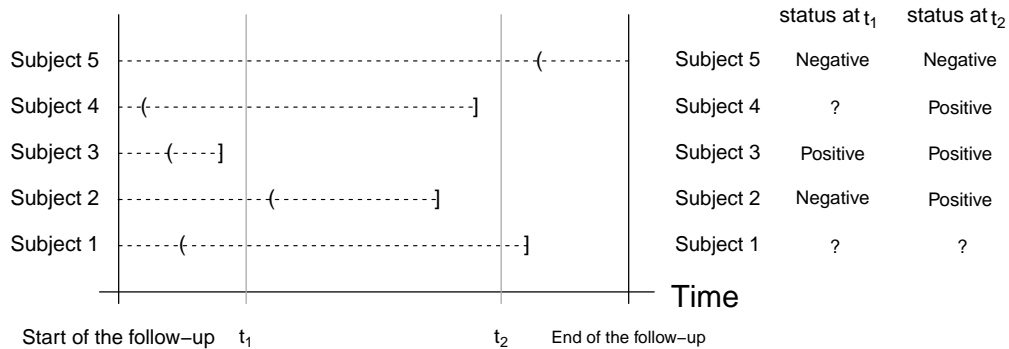


Figure 2: Example of the status of five interval censored individuals at two different times t_1 and t_2 . Left bracket (stands for the last revision time before the event happened and the right square bracket] represents the next revision time when the event has already taken place.

3.2 Interval censorship

Interval censorship arises when the event times are not exactly observed and it is only known that they lie in a certain range. The final available information consists of intervals $(L, R]$ with $L \leq R$ and $T \in (L, R]$, where T represents the time-to-event event variable; L stands for the last observable time before the event occurs and R indicates the revision time when the event was first observed. We will use the notation of left-open and right-closed intervals although the edges L and R may take 0 and ∞ values, respectively or may represent an exact observation ($L = R$).

An example of the status of some individuals under an interval censoring scheme is shown in Figure 2. Once fixed a time t , positive individuals are those whose observable interval meets $R \leq t$ (subject 3 at t_1 and t_2 and subjects 2 and 4 at t_2). Negative individuals are those for who $t < L$ (subject 5 at t_1 and t_2 and subject 2 at t_1). Finally, subjects for who $L \leq t < R$ are going to be referred as *undefined individuals* because no affirmation can be made about their status at t .

References regarding C/D ROC curve estimation methods for interval-censored data are not as numerous as they are for the right censoring context. Li and Ma [2011] proposed to remove the undefined individuals from the original sample (naive procedure). Corre-

sponding cumulative sensitivity and dynamic specificity are subsequently estimated with the remaining individuals. The bias in the estimation is quantified in terms of the undefined subjects proportion. The authors state that as much as this proportion increases, the accuracy of the estimator decreases. Jacqmin-Gadda et al. [2016] suggested additional approaches for particular frameworks. Wu et al. [2018] propose a non-parametric estimator of the joint distribution function of the bivariate vector (*Marker*, *Time-to-event*) and the corresponding marginal distribution of the *Marker*. On simulation studies, as the authors pointed out, the bias increases with the correlation between the marker and the event time.

4 The Mixed subjects C/D ROC curve estimator

4.1 General context

The criterion used by Martínez-Cambor et al. [2016] to estimate the C/D ROC curve under right censorship (also used later by Li et al. [2018]) can be extended to more general situations. Once set a time point t , the underlying procedure accounts censored observations before t as partially both, positive and negative. The information brought by these individuals is used to get adequate estimators of the real probabilities of being positive (respectively, negative). Based on those estimates, censored individuals are partially allocated to the positive and negative groups.

Notice that, for any random variable \mathbf{Y} , once fixed the time t , the cumulative sensitivity can be expressed as

$$\begin{aligned} Se_t^C(c) &= \mathcal{P}(X > c \mid T \leq t) \\ &= \frac{\mathcal{P}(X > c, T \leq t)}{\mathcal{P}(T \leq t)} \\ &= \frac{\mathbb{E}_{\mathbf{Y}}[\mathcal{P}(X > c, T \leq t \mid \mathbf{Y})]}{\mathbb{E}_{\mathbf{Y}}[\mathcal{P}(T \leq t \mid \mathbf{Y})]}, \end{aligned}$$

where T and X stand, as usual, for the time-to-event and marker variables, respectively and $c \in \mathbb{R}$. Assuming that \mathbf{Y} is a random variable containing all the information available in the sample (marker values, X , and other relevant information, \mathbf{W}), we have

$$Se_t^C(c) = \frac{\int \mathcal{P}(T \leq t \mid \mathbf{W} = \mathbf{w}, X = x) I_{(c, \infty)}(x) dH(\mathbf{w}) dx}{\int \mathcal{P}(T \leq t \mid \mathbf{W} = \mathbf{w}, X = x) dH(\mathbf{w}) dx},$$

where $H(\cdot)$ stands for the CDF of the variable \mathbf{W} and I is the indicator function ($I_A(x)$ takes the value 1 if $x \in A$ and 0 otherwise). Analogously, at time t , the dynamic specificity can be written as

$$\begin{aligned} Sp_t^{\mathbb{D}}(c) &= \frac{\mathbb{E}_{\mathbf{Y}}[\mathcal{P}(X \leq c, T > t \mid \mathbf{Y})]}{\mathbb{E}_{\mathbf{Y}}[\mathcal{P}(T > t \mid \mathbf{Y})]} \\ &= \frac{\int \mathcal{P}(T > t \mid \mathbf{W} = \mathbf{w}, X = x) I_{(-\infty, c]}(x) dH(\mathbf{w}) dx}{\int \mathcal{P}(T > t \mid \mathbf{W} = \mathbf{w}, X = x) dH(\mathbf{w}) dx}. \end{aligned}$$

Let $\{y_i\}_{i=1}^n$ be a random sample, independent and identically distributed where each y_i contains all the available information for the i th subject, $i \in \{1, \dots, n\}$ (again, the marker value x_i is included in each y_i as well as other relevant information, w_i). The natural semi-parametric estimators for $Se_t^{\mathbb{C}}$ and $Sp_t^{\mathbb{D}}$ are

$$\begin{aligned} \hat{Se}_t^{\mathbb{C}}(c) &= \frac{\sum_{i=1}^n I_{(c, \infty)}(x_i) \cdot \hat{p}_i(t)}{\sum_{i=1}^n \hat{p}_i(t)} \quad \text{and} \\ \hat{Sp}_t^{\mathbb{D}}(c) &= \frac{\sum_{i=1}^n I_{(-\infty, c]}(x_i) \cdot [1 - \hat{p}_i(t)]}{\sum_{i=1}^n [1 - \hat{p}_i(t)]}, \end{aligned}$$

respectively and, for $i \in \{1, \dots, n\}$, $\hat{p}_i(t)$ ($= \hat{p}_{i,n}(t)$) are adequate estimators for $p_i(t) = \mathcal{P}(t_i \leq t \mid y_i)$. Of course, when t_i is observed, $\hat{p}_i(t)$ should take the value 1, if $t_i \leq t$ and 0 when $t_i > t$. Directly, the Mixed subjects C/D ROC curve estimator associated to a time point t , is defined by

$$\hat{\mathcal{R}}_t^{\mathbb{C}/\mathbb{D}}(p) = \hat{Se}_t^{\mathbb{C}}([1 - \hat{Sp}_t^{\mathbb{D}}]^{-1}(p)), \quad p \in [0, 1],$$

where $[1 - \hat{Sp}_t^{\mathbb{D}}]^{-1}(\cdot) = \inf\{x : 1 - \hat{Sp}_t^{\mathbb{D}}(x) \leq \cdot\}$.

4.2 The right censorship context

Martínez-Camblor et al. [2016] proposed two different estimation procedures for the probabilities $p_i(t)$ ($1 \leq i \leq n$) in the right-censored case. Notice that, in this context, each sample observation is in the way $y_i = \{z_i, \delta_i, x_i\}$ where, for $i \in \{1, \dots, n\}$, z_i depicts the observed time, δ_i stands for the status of the observation and x_i for the marker value.

In the first procedure, the estimators are given by $\hat{p}_i(t) = 1 - \left(\hat{\mathcal{S}}_C(t|x_i) / \hat{\mathcal{S}}_C(z_i|x_i) \right)$, $1 \leq i \leq n$, where $\hat{\mathcal{S}}_C(\cdot)$ is the approximation for the survival function obtained from a Cox proportional hazards model. With the second one, the estimators are given by $\hat{p}_i(t) = 1 - \left(\hat{\mathcal{S}}_{KM}(t) / \hat{\mathcal{S}}_{KM}(z_i) \right)$, $1 \leq i \leq n$, being $\hat{\mathcal{S}}_{KM}(\cdot)$ the Kaplan-Meier estimator

for the survival function. Li et al. [2018] expressed the probabilities $p_i(t)$ ($1 \leq i \leq n$) in terms of the conditional survival distribution of the time-to-event variable given the marker. The authors suggested the use of the kernel weighted Kaplan-Meier estimator for approximating this survival function. Formally, the expressions for these estimators are $\hat{p}_i = 1 - \left(\hat{\mathcal{S}}_L(t|x_i)/\hat{\mathcal{S}}_L(z_i|x_i) \right)$, ($1 \leq i \leq n$) where $\hat{\mathcal{S}}_L(\cdot)$ is the survival function estimator conditioned to the marker value:

$$\hat{\mathcal{S}}_L(t|x_i) = \prod_{\{s \in \Omega, s \leq t\}} \left\{ 1 - \frac{\sum_j K_h(x_j, x_i) I(z_j = s) \delta_j}{\sum_j K_h(x_j, x_i) I(z_j \geq s)} \right\};$$

Ω stands for the set of distinct event times; I is the indicator function and $K_h(\cdot, \cdot)$ stands for the uniform kernel. An important handicap is the impact of the bandwidth parameter associated to kernel procedures on the obtained estimations.

4.3 The interval censorship context

Under the interval censored scheme previously described in Section 3, each sample observation is in the way $y_i = \{(l_i, r_i, x_i)\}$, where for $i \in \{1, \dots, n\}$, l_i and r_i are the lower and upper edges, respectively, of the observed interval containing the unobserved event time t_i ($l_i \leq t_i \leq r_i$) and x_i stands for the marker value. It is assumed the common condition of non-informative censoring which guarantees that the mechanism generating censorship is independent of the underlying distribution of the time-to-event variable T . It is also assumed that this mechanism is independent of the considered marker X . Once set the point time t , we propose the following estimators for the probabilities $p_i(t)$,

$$\hat{p}_i(t) = 1 - \frac{\hat{\mathcal{S}}_n(t|x_i) - \hat{\mathcal{S}}_n(v_i|x_i)}{\hat{\mathcal{S}}_n(u_i|x_i) - \hat{\mathcal{S}}_n(v_i|x_i)},$$

where $u_i = \min\{l_i, t\}$ and $v_i = \max\{r_i, t\}$ ($1 \leq i \leq n$). We suppose as well that $\mathcal{S}(t|X)$ is specified by a proportional hazards model $\mathcal{S}(t|X) = \mathcal{S}_0(t)e^{X'\beta}$, where $\mathcal{S}_0(t)$ is the baseline survivor function and X' the transpose of the marker variable. We propose the use of the semi-parametric proportional hazards model under interval censorship given by Finkelstein [1986] for estimating the survival function $\mathcal{S}(\cdot)$. The resulting estimator is a step function outside the Turnbull's intervals [Turnbull, 1976] and remains undefined inside them, because one range of solutions are possible. Among them, we have chosen the linear interpolation between the upper edge of one step and the lower of the next one. Obviously, as it happened

in the right censoring context, the resulting $\hat{p}_i(t)$ estimations will take the value 1 for positive subjects and 0 for the negative ones.

Under the aforementioned assumptions, the obtained approximation, $\hat{\mathcal{S}}_n(\cdot)$ is a consistent estimator of $\mathcal{S}(\cdot)$ and then, the respective $\hat{p}_i(t)$ are consistent estimators for $p_i(t)$ ($1 \leq i \leq n$). The full derivation of the uniform strong consistence of the proposed *Mixed subjects* estimator for the C/D ROC curve under interval censorship is provided as appendix.

5 Computational considerations

Most of the statistical commercial software includes procedures for computing the ROC curve and of course R , free-available at CRAN (www.r-project.org), does. Although one of the most popular R packages is the pROC [Robin et al., 2011], there exists a considerable number of packages dealing with ROC curve estimation and other related topics, such as areas under the curve, partial areas or confidence intervals for the AUC, sensitivity and specificity. In a non-exhaustive list we include: fbroc, ROC632, rocc, roccv, rocNIT, ROCS, auRoc, ggROC or nproc.

Computing the C/D ROC curve and/or the corresponding areas under the curve has become, as well, a routine task in the prognosis procedures. Available R packages where most of the estimation procedures mentioned in Section 3 have been implemented are: survivalROC, survAUC, timeROC and smoothROCtime, among others.

The *Mixed subjects* approach for the estimation of the C/D ROC curve in the right-censoring case has also been implemented in R packages: the nsROC and the tdROC. The former provides functions for computing both, Martínez-Cambolor et al. [2016] and Li et al. [2018] methods, together with another non-standard procedures to estimate different ROC curve elements. The latter implements the Li et al. [2018] method.

Regarding the estimation of the C/D ROC curve under interval censorship, the package intcensROC is currently available and gives approximations for the C/D ROC curve and the corresponding AUC according to the Wu et al. [2018] proposal. We have also developed the R function intCDroc where the *Mixed subjects* estimator of the C/D ROC curve under interval censorship is implemented. The corresponding survival function estimates on which the *Mixed subjects* estimator is based are computed using ic-sp function from icenReg

package [Anderson-Bergman, 2017]. The code of the function `intCDroc` is provided as appendix.

6 Simulation results

To illustrate the performance of the proposed *Mixed subjects* estimator we have computed Monte Carlo simulations under two different scenarios. The following situations for the time-to event variable, T , were considered:

- **Scenario I:** it is assumed that the event times come from a proportional hazards model, $\mathcal{S}(t|X) = \mathcal{S}_0(t)e^{X' \cdot \beta}$. $\mathcal{S}_0(t)$ is generated from an exponential distribution of mean 2. The marker variable X has been generated from a normal distribution with mean 0 and variance 1. For the parameter β , we have considered three different values: $\beta = 1.250$, $\beta = 0.726$ and $\beta = 0.390$ in order to represent a weak, medium and strong association respectively, between the marker and the time-to-event variable.
- **Scenario II:** event times are generated from a non-proportional hazards model, particularly, $\log(T) = \rho \cdot X + \sqrt{1 - \rho^2} \cdot M$, where X is the marker and M is an independent random variable, both following a normal distribution with mean 0 and variance 1. Like in the previous scenario, we have taken three values for ρ to study different correlation grades between the marker and time-to-event variables ($\rho = -0.730$, $\rho = -0.536$ and $\rho = -0.327$).

C/D ROC curves are considered at times corresponding to the first quartile, Q_1 , second quartile, Q_2 and third quartile, Q_3 of the time-to-event distribution, T , in each scenario. The modeled strong, medium and weak relations between the marker values and the event times give rise to areas under the real C/D ROC curves at point Q_2 of 0.85, 0.75 and 0.65. Real C/D ROC curves at the considered times are plotted in Figures 3 and 4 for the scenario I and II, respectively.

The interval censored mechanism has been generated as indicated in Gómez et al. [2009], to meet the non-informative censoring condition: the lower and upper edges of each interval satisfy, respectively, $L_i = \max\{T_i - U_i^{(1)}, T_i + U_i^{(2)} - a\}$ and $R_i = \min\{T_i + U_i^{(2)}, T_i - U_i^{(1)} + a\}$, for $i \in \{1, \dots, n\}$, where $U^{(1)}$ and $U^{(2)}$ are independent continuous variables uniformly distributed in the interval $(0, a)$, with $a = 1/2$. Besides, 35% of intervals have been forced

to be right-open ($R = \infty$). In order to compare the results obtained by our proposal with the Li and Ma [2011] method, this estimator has also been computed and used as reference. The results for the proportional hazards scenario are shown in Table 1 and those for the non-proportional hazards one in Table 2. The means and standard deviations of the estimated AUCs are reported for both estimators, together with the means and standard deviations for the integrate of the absolute error (IAE): $\int_0^1 |\hat{\mathcal{R}}_t^{\text{C/D}}(p) - \mathcal{R}_t^{\text{C/D}}(p)| dp$, where $\mathcal{R}_t^{\text{C/D}}$ is the true C/D ROC curve and $\hat{\mathcal{R}}_t^{\text{C/D}}$ its estimator. It has been computed from 2000 Monte Carlo simulations with sample sizes $n = 100$, $n = 250$ and $n = 500$. The percentage of undefined individuals is shown in each situation considered .

Given the obtained results, we highlight that the *Mixed subjects* estimator shows to have smaller IAE in all considered cases. The Li and Ma [2011] estimator tends to over-estimate the AUC, as well, in all situations. The IAEs are smaller for the C/D ROC estimates obtained by the two procedures at time $t = Q_3$, where the censoring percentage is larger than in the other cases. Obviously, these IAEs are greater, for both estimators, the weaker is the association between the marker and time-to-event variables and the smallest ones arise for the bigger sample size ($n = 500$). Li and Ma [2011] approximation shows to have a slightly less IAE in the non-proportional hazards scenario than in the proportional hazards one. These IAEs for the *Mixed subjects* estimator are similar in both scenarios.

7 Real world application

The performance of the proposed estimator is illustrated on a real-world problem. In particular, we consider *the Hepatitis C data set*, which has motivated this research. The Hepatitis C virus (HCV) is a major cause of chronic hepatitis, cirrhosis, and hepatocellular carcinoma (HCC), affecting over 170 million people world-wide [Lavanchy, 2009, Hoofnagle, 2002]. HCV does not integrate its genetic material into the host genome so it requires continuous replications to maintain chronic infection. As a consequence, HCV may promote carcinogenesis through chronic inflammation of hepatic cells, with the subsequent fibrosis [Matsuzaki and Murata, 2007]. The most severe forms of fibrosis yield to cirrhosis and in a long term, to the development of HCC. An early prognosis of the severity of the fibrosis allows physicians to identify individuals at risk who may receive appropriate treatments to prevent the HCC. The main goal of this study is to assess the predictive ability of certain

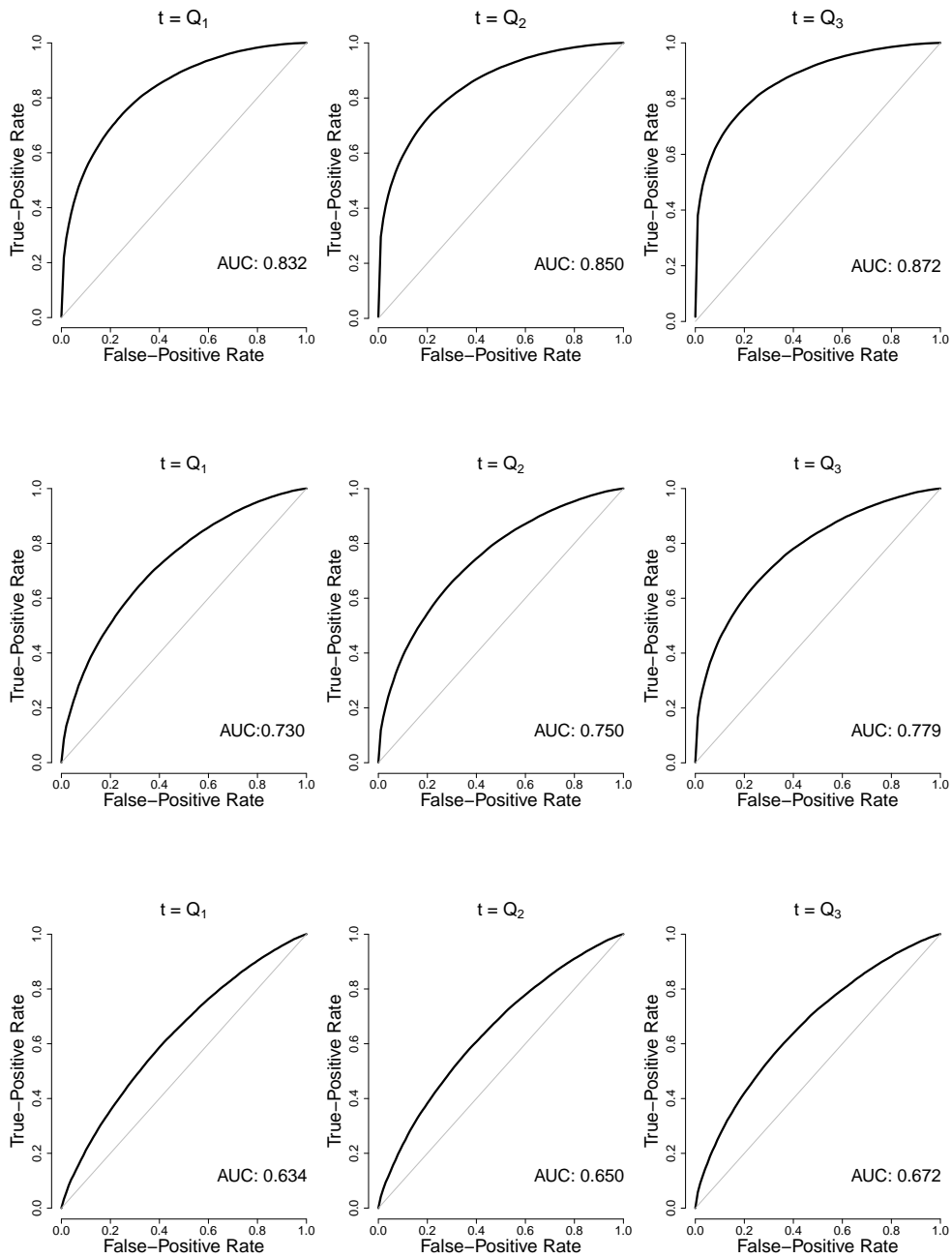


Figure 3: **Scenario I** - Real C/D ROC curves under the scenario of time events from a proportional hazards model. Top: strong association between biomarker and time-to-event variables ($AUC \sim 0.85$); middle: medium association ($AUC \sim 0.75$) and bottom: weak association ($AUC \sim 0.65$).

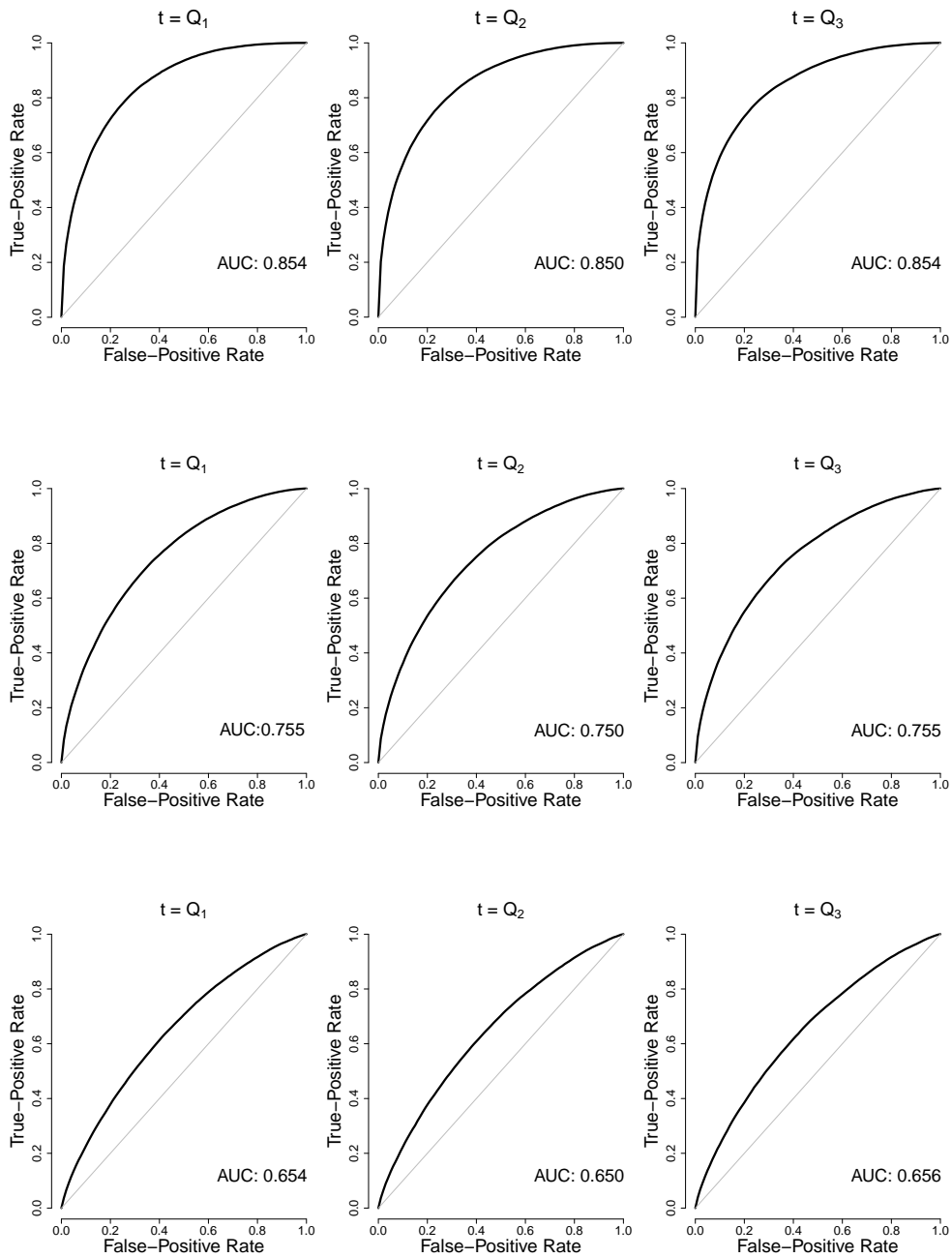


Figure 4: **Scenario II** - Real C/D ROC curves under the scenario of event times coming from a non-proportional hazards model. Top: strong correlation between biomarker and time-to-event variables ($AUC \sim 0.85$); middle: medium correlation ($AUC \sim 0.75$) and bottom: weak correlation ($AUC \sim 0.65$).

n	t	$\mathcal{A}^{\mathbb{C}/\mathbb{D}}(t)$	% U.	<i>Mixed subjects</i>		<i>Li&Ma</i>	
				<i>AUC</i>	<i>Error</i>	<i>AUC</i>	<i>Error</i>
100	Q_1	0.832	21	0.836 ± 0.05	0.049 ± 0.02	0.857 ± 0.06	0.066 ± 0.03
100	Q_2	0.850	23	0.847 ± 0.04	0.041 ± 0.02	0.864 ± 0.04	0.049 ± 0.02
100	Q_3	0.872	26	0.863 ± 0.04	0.038 ± 0.02	0.879 ± 0.04	0.045 ± 0.02
250	Q_1	0.832	21	0.837 ± 0.03	0.031 ± 0.01	0.860 ± 0.03	0.045 ± 0.02
250	Q_2	0.850	22	0.846 ± 0.02	0.026 ± 0.01	0.867 ± 0.03	0.034 ± 0.01
250	Q_3	0.872	25	0.864 ± 0.02	0.024 ± 0.01	0.882 ± 0.03	0.029 ± 0.01
500	Q_1	0.832	21	0.833 ± 0.02	0.022 ± 0.01	0.857 ± 0.02	0.034 ± 0.01
500	Q_2	0.850	23	0.833 ± 0.02	0.019 ± 0.01	0.866 ± 0.02	0.026 ± 0.01
500	Q_3	0.872	25	0.863 ± 0.02	0.018 ± 0.01	0.882 ± 0.02	0.022 ± 0.01
100	Q_1	0.730	21	0.729 ± 0.06	0.061 ± 0.03	0.748 ± 0.07	0.075 ± 0.03
100	Q_2	0.750	24	0.744 ± 0.05	0.053 ± 0.02	0.768 ± 0.06	0.062 ± 0.03
100	Q_3	0.779	27	0.762 ± 0.05	0.050 ± 0.02	0.788 ± 0.05	0.057 ± 0.02
250	Q_1	0.730	21	0.725 ± 0.03	0.038 ± 0.02	0.746 ± 0.04	0.048 ± 0.02
250	Q_2	0.750	23	0.742 ± 0.03	0.034 ± 0.02	0.765 ± 0.03	0.040 ± 0.02
250	Q_3	0.779	26	0.763 ± 0.03	0.034 ± 0.02	0.790 ± 0.03	0.038 ± 0.02
500	Q_1	0.730	21	0.728 ± 0.02	0.026 ± 0.01	0.750 ± 0.03	0.037 ± 0.02
500	Q_2	0.750	24	0.743 ± 0.02	0.023 ± 0.01	0.768 ± 0.02	0.031 ± 0.01
500	Q_3	0.779	27	0.765 ± 0.02	0.025 ± 0.01	0.792 ± 0.02	0.028 ± 0.01
100	Q_1	0.634	21	0.635 ± 0.06	0.067 ± 0.07	0.648 ± 0.07	0.080 ± 0.04
100	Q_2	0.650	24	0.643 ± 0.06	0.058 ± 0.06	0.661 ± 0.06	0.067 ± 0.03
100	Q_3	0.672	28	0.657 ± 0.05	0.056 ± 0.06	0.680 ± 0.06	0.064 ± 0.03
250	Q_1	0.634	21	0.631 ± 0.04	0.041 ± 0.02	0.645 ± 0.05	0.050 ± 0.02
250	Q_2	0.650	24	0.642 ± 0.04	0.038 ± 0.02	0.662 ± 0.04	0.044 ± 0.02
250	Q_3	0.672	27	0.638 ± 0.04	0.039 ± 0.02	0.682 ± 0.04	0.044 ± 0.02
500	Q_1	0.634	21	0.632 ± 0.03	0.029 ± 0.01	0.647 ± 0.03	0.037 ± 0.02
500	Q_2	0.650	24	0.644 ± 0.02	0.026 ± 0.01	0.664 ± 0.03	0.032 ± 0.01
500	Q_3	0.672	27	0.660 ± 0.02	0.027 ± 0.01	0.686 ± 0.03	0.022 ± 0.01

Table 1: **Scenario I**, Proportional hazards model. Mean \pm standard deviation of AUC and $\int_0^1 |\hat{\mathcal{R}}_t^{\mathbb{C}/\mathbb{D}}(p) - \mathcal{R}_t^{\mathbb{C}/\mathbb{D}}(p)| dp$, where $\mathcal{R}_t^{\mathbb{C}/\mathbb{D}}$ is the true C/D ROC curve and $\hat{\mathcal{R}}_t^{\mathbb{C}/\mathbb{D}}$ is its estimator, computed from 2000 Monte Carlo simulations with sample size $n = 100$, $n = 250$ and $n = 500$. $\mathcal{A}^{\mathbb{C}/\mathbb{D}}(t)$ is $\int_0^1 \mathcal{R}_t^{\mathbb{C}/\mathbb{D}}(p) dp$ and %U depicts the expected percentage of undefined individuals.

n	t	$\mathcal{A}^{\text{C}/\text{D}}(t)$	%U.	<i>Mixed subjects</i>		<i>Li&Ma</i>	
				<i>AUC</i>	<i>Error</i>	<i>AUC</i>	<i>Error</i>
100	Q_1	0.854	27	0.847 ± 0.04	0.045 ± 0.02	0.897 ± 0.05	0.067 ± 0.03
100	Q_2	0.850	29	0.841 ± 0.04	0.043 ± 0.02	0.882 ± 0.04	0.054 ± 0.02
100	Q_3	0.854	30	0.843 ± 0.04	0.042 ± 0.02	0.872 ± 0.04	0.048 ± 0.02
250	Q_1	0.854	27	0.846 ± 0.03	0.028 ± 0.01	0.899 ± 0.03	0.052 ± 0.02
250	Q_2	0.850	29	0.843 ± 0.02	0.026 ± 0.01	0.883 ± 0.02	0.041 ± 0.02
250	Q_3	0.854	30	0.846 ± 0.02	0.026 ± 0.01	0.878 ± 0.03	0.035 ± 0.02
500	Q_1	0.854	27	0.847 ± 0.02	0.020 ± 0.01	0.901 ± 0.02	0.049 ± 0.02
500	Q_2	0.850	29	0.842 ± 0.02	0.019 ± 0.01	0.884 ± 0.02	0.038 ± 0.01
500	Q_3	0.854	30	0.844 ± 0.02	0.020 ± 0.01	0.877 ± 0.02	0.029 ± 0.01
100	Q_1	0.755	27	0.747 ± 0.06	0.058 ± 0.03	0.795 ± 0.07	0.081 ± 0.04
100	Q_2	0.750	29	0.738 ± 0.05	0.055 ± 0.03	0.778 ± 0.06	0.066 ± 0.03
100	Q_3	0.755	30	0.740 ± 0.05	0.055 ± 0.03	0.771 ± 0.06	0.062 ± 0.03
250	Q_1	0.755	26	0.748 ± 0.03	0.036 ± 0.02	0.797 ± 0.04	0.058 ± 0.03
250	Q_2	0.750	28	0.741 ± 0.03	0.033 ± 0.02	0.782 ± 0.03	0.047 ± 0.02
250	Q_3	0.755	30	0.743 ± 0.03	0.034 ± 0.02	0.776 ± 0.04	0.043 ± 0.02
500	Q_1	0.755	27	0.748 ± 0.02	0.026 ± 0.01	0.800 ± 0.03	0.051 ± 0.02
500	Q_2	0.750	28	0.741 ± 0.02	0.024 ± 0.01	0.782 ± 0.02	0.040 ± 0.02
500	Q_3	0.755	30	0.743 ± 0.02	0.025 ± 0.01	0.776 ± 0.02	0.033 ± 0.01
100	Q_1	0.634	26	0.649 ± 0.06	0.066 ± 0.03	0.686 ± 0.08	0.089 ± 0.04
100	Q_2	0.650	29	0.640 ± 0.06	0.059 ± 0.03	0.668 ± 0.07	0.072 ± 0.03
100	Q_3	0.672	30	0.641 ± 0.06	0.059 ± 0.03	0.663 ± 0.07	0.068 ± 0.03
250	Q_1	0.634	26	0.649 ± 0.04	0.041 ± 0.02	0.683 ± 0.05	0.060 ± 0.03
250	Q_2	0.650	29	0.644 ± 0.04	0.037 ± 0.02	0.672 ± 0.04	0.050 ± 0.02
250	Q_3	0.672	30	0.644 ± 0.04	0.037 ± 0.02	0.668 ± 0.04	0.046 ± 0.02
500	Q_1	0.654	27	0.650 ± 0.03	0.029 ± 0.01	0.686 ± 0.04	0.047 ± 0.02
500	Q_2	0.650	28	0.644 ± 0.02	0.026 ± 0.01	0.673 ± 0.03	0.037 ± 0.02
500	Q_3	0.656	30	0.645 ± 0.02	0.026 ± 0.01	0.669 ± 0.03	0.033 ± 0.01

Table 2: **Scenario II**, Non-proportional hazard model scenario. Mean \pm standard deviation of AUC and $\int_0^1 |\hat{\mathcal{R}}_t^{\text{C}/\text{D}}(p) - \mathcal{R}_t^{\text{C}/\text{D}}(p)| dp$, where $\mathcal{R}_t^{\text{C}/\text{D}}$ is the true C/D ROC curve and $\hat{\mathcal{R}}_t^{\text{C}/\text{D}}$ is its estimator, computed from 2000 Monte Carlo simulations with sample size $n = 100$, $n = 250$ and $n = 500$. $\mathcal{A}^{\text{C}/\text{D}}(t)$ is $\int_0^1 \mathcal{R}_t^{\text{C}/\text{D}}(p) dp$ and %C depicts the expected percentage of undefined individuals.

polymorphism combined with clinical variables to determine the worsening in the fibrosis stage, in individuals infected by the HCV virus.

A total of 722 patients diagnosed with chronic HCV infection who underwent medical revision at centers University Central Hospital of Asturias (HUCA) in Oviedo (Asturies, Spain), La Princesa Hospital in Madrid (Spain) and San Agustin Hospital in Avilés (Asturies, Spain), were enrolled in a retrospective study, during the period between January and December of 2016. The fibrosis stage of the patients and several polymorphisms expressions were registered. In particular, we focus in variant *rs27044* of the *ERAP1*, whose analyzed combination alleles were *C/C*, *C/G* and *G/G*. The *HCV infection diagnostic date* and other background variables such us *age* and *gender* were also collected. The fibrosis stages values range from 0 (free of fibrosis) to 5 (worst condition) and were categorized as a binary outcome. The highest stages (values 4 and 5) were considered as a worsening in the disease. The remainder values depicted not having or suffer the disease in a low stage. A total of 154 patients were finally removed from the original sample because of the lack of HCV infection diagnosis date.

In this time-to-event study, the event of interest for an individual is to have a worsening in the disease. The period that elapses since the HCV infection diagnosis date until that event occurs, represents the time-to-event variable. Two different situations may arise: at the scheduled revision the subject does not have fibrosis or suffers it in its lowest stages or, at that revision, it is found that s/he presents a highest stage of fibrosis. However, in the later, physicians do not know the exact moment of the onset of that stage. The only information available is the range between the HCV infection diagnosis date, when the subject was event-free, and the revision date, when the fibrosis status is determined.

It is well-known that the age is a risk factor for the worsening of the fibrosis stage in patients. We are going to consider the continuous marker defined as a linear combination of the *age* and the three categories of the *rs27044* variant of *ERAP1*. The weights for the linear combination will be the coefficients from the correspondent proportional hazards regression model under interval censorship. The discriminatory ability of the marker for predicting the potential worsening in the fibrosis disease status will be evaluated at four different times: 5, 10, 15 and 20 years.

The time-free of worsening in the fibrosis status is shown in Figure 5. During the two first years of the follow-up, approximately 40% of patients have suffered a deterioration in

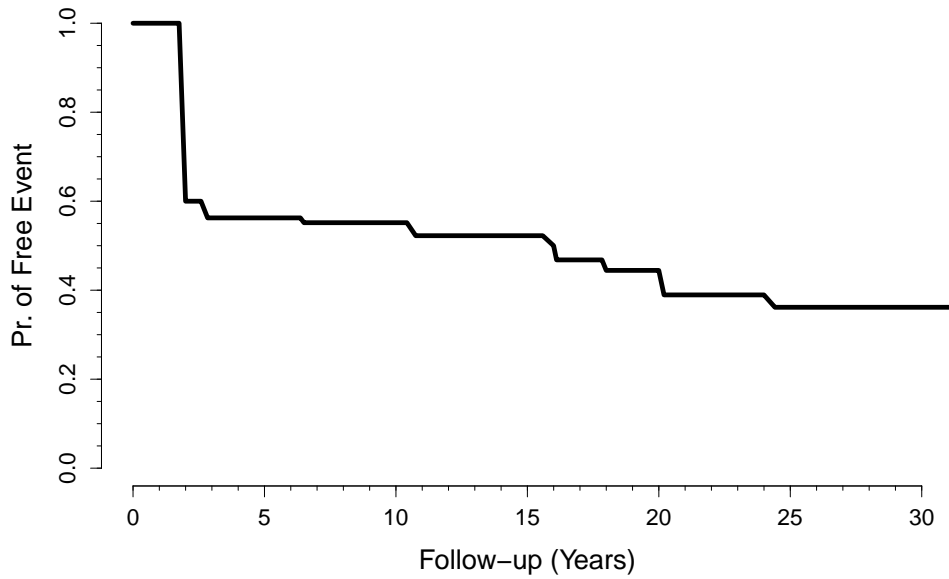


Figure 5: Turnbull estimation for the survival function on all individuals included in the study, except those for who the fibrosis diagnosis date has been missed. Linear interpolation was made in the Turnbull intervals.

their fibrosis status. After that time, the worsening takes place smoothly and over 25 years of the follow-up, around the 36% remain in the lowest stages of the disease. The median of the time until getting more advanced stages in fibrosis is 16 years. The mean of the length of the finite observed intervals is around 20 years with a standard deviation of 7 years. The shortest and highest ones have a length of 2 and 37 years, respectively.

Table 3 shows the number of positive individuals, the number of negative and the number of those whose status can not be determined with the available information, at the four considered points of time: 5, 10, 15 and 20 years. It can be observed that there is a large number of undefined subjects for every set of time and that this number does not increase as long as we consider greatest times.

In order to assess the predictive ability of the marker, both estimations of C/D ROC curve under interval censorship (*Mixed subjects* and Li and Ma [2011]) have been computed at times: 5, 10, 15 and 20 years. The graphics in Figure 6 show that Li and Ma [2011] estimator provides higher values for the AUCs at all considered points of time. However, as it was seen in the Monte Carlo simulations, this estimation method tends to over estimate

	t = 5	t = 10	t = 15	t = 20
Positive	10	41	90	148
Negative	188	148	107	71
Undefined subjects	290	299	291	269

Table 3: Number of Positive, Negative and Undefined individuals at the considered points of time. Individuals with *age* missing values were not taking into account.

the AUC. These values, computed with the *Mixed subjects* estimator, increase as we consider points further away in time. Particularly, the AUC remains quite similar at times $t = 5$ and $t = 10$, and then increases progressively, reaching the higher values at time $t = 20$. On the other hand, for the Li and Ma [2011] estimator, AUCs are bigger at $t = 5$ and then, decrease progressively until they take their lowest value at $t = 20$.

Although more optimistic results are achieved with the Li and Ma [2011] proposal, it dispenses with a lot of information from the sample. This behaviour matches with that shown in Monte Carlo simulations. It is worth to notice that if we want to estimate the C/D ROC curve at a point t and we just know the status of the subjects at times $t_1 < t < t_2$, the estimator for the cumulative sensitivity, for example, at a fixed threshold c , is given by the probability that the marker takes values greater than the threshold, conditioned to the event time were less or equal than t_1 instead of less or equal than the considered point time t . A more realistic assessment of the marker has been achieved with the *Mixed subjects* estimator.

8 Discussion

Interval censored data arise frequently in prognosis studies with time-to-event outcomes. Subjects are not constantly monitored and at the revision times is only known whether or not they experienced the event in study. Besides, even in the first case, it is not possible to determine exactly when the event took place.

When estimating the C/D ROC curve, a common practice dealing with interval-censored data is to impute the unknown event time to some point in the observed interval and then to use any of the C/D ROC curve estimators for the right-censored case. To the best

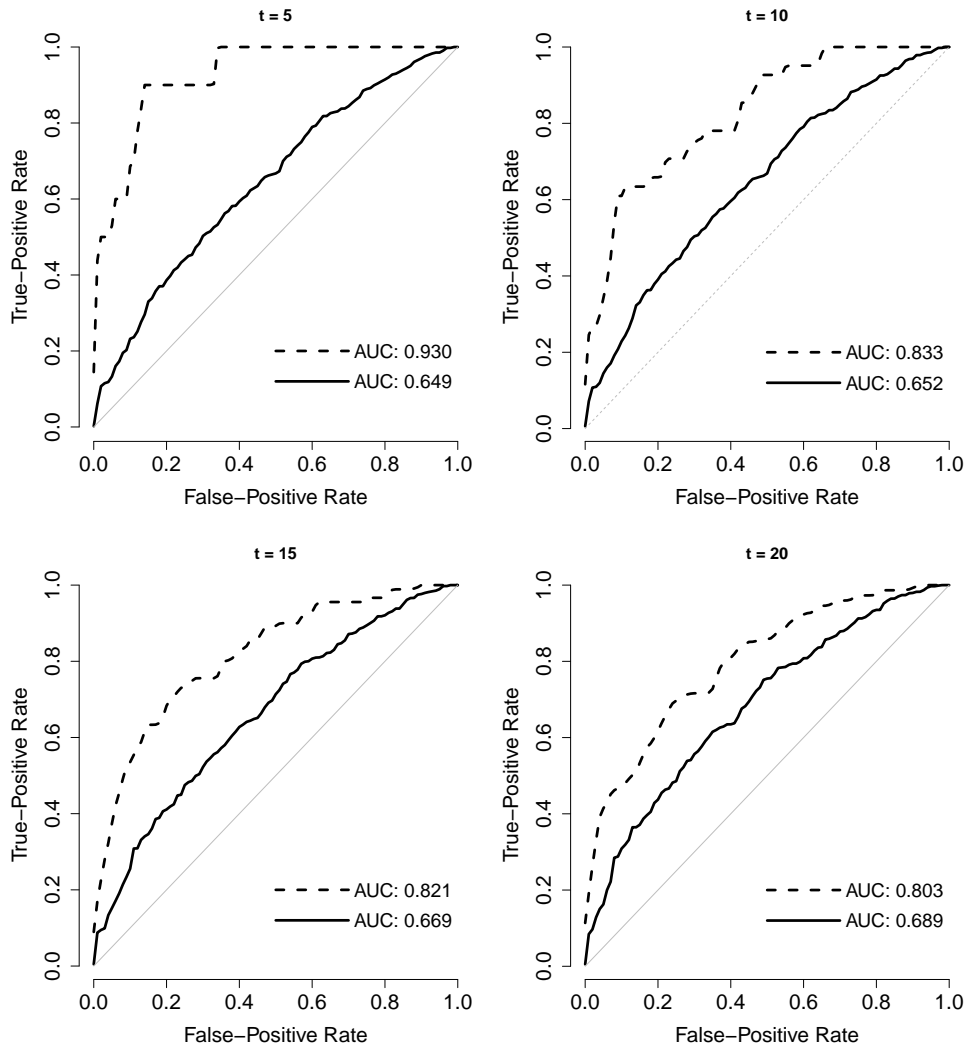


Figure 6: Left: C/D ROC curve estimations and their corresponding AUCs for marker X at different years: $t = 5$, upper left; $t = 10$, upper right; $t = 15$, bottom left, $t = 20$ bottom right. The dotted lines are those obtained with the Li and Ma [2011] procedure and the solid lines represent those computed with the *Mixed subjects* estimator.

of our knowledge, only an estimation method for the C/D ROC curve under the general context of interval censorship has been published. Once set the time t , those subjects whose observed interval contains that point are removed from the sample and the approximation is made with the remaining individuals. Clearly, their status as positive or negative is unambiguously defined.

In this paper we propose a new estimator for the C/D ROC curve for interval-censored data, where the information provided for all individuals in the sample is used. It is an extension to the interval-censored scenario of the *Mixed subjects* approach, already proposed by Martínez-Cambor et al. [2016] in the right-censored case. Individuals whose status is unknown at a time t are partially allocated to both, positive and negative groups, with certain probabilities. These probabilities are estimated from different approaches according to the censoring scenario. In particular, our proposal is based on the proportional hazards model for interval-censored data. Besides, we prove the uniform consistency of the estimator. This result can be easily applied to the previously proposed procedures in the right censoring context, just treated empirically by the mentioned authors.

In the Monte Carlo study conducted under two different scenarios (event times from a proportional and a non-proportional hazards models, respectively), the *Mixed subjects* estimator has shown to have less integrate absolute error (IAE) than the already existing method, which tends, as well, to over estimate the corresponding areas under the C/D ROC curve.

Our proposal was applied to the Hepatitis C data set. We defined a continuous marker as the linear combination of the age and the scores from a proportional hazards model with variant *rs27044* of the polymorphism *ERAP1* as covariate. Our goal was to assess the ability of this marker to identify individuals at risk of reaching advanced stages of fibrosis over time (and the subsequent development of carcinoma). C/D ROC curve approximations were obtained with both, *Mixed subjects* and Li and Ma [2011] estimators, at four different times: 5, 10, 15 and 20 years. The former shows to be more realistic than the latter, which dispenses with more than a half of the available sample observations at all considered times.

In short, the proposed *Mixed subjects* estimator proves to have good asymptotic properties and improves the existing procedures, even when the underlying predictive model is misspecified (scenario II). The methodology is quite general and could be applied to other types of data and censorship patterns. Notice that its practical implementation requires

a linear interpolation, whose impact on the final estimation could be the source of future researches.

Funding

This research was supported by the grant FC-GRUPIN- IDI/2018/000132 of the Asturias Government. PM-C is supported by the Spanish Ministry of Economy, Industry and Competitiveness; State Research Agency; and FEDER funds - MTM2017-89422-P.

Appendix A: Strong uniform consistency

Next theorem guarantees the strong uniform consistency for the Mixed subject C/D ROC curve estimator on the interval censorship context. The proof is more general and useful when the probabilities computed for the so-called undefined subjects are estimated from a consistent predictive model.

Theorem. Let $\{l_i, r_i, x_i\}_{i=1}^n$ be an independent random sample where for the i th subject x_i stands for its marker value and it is known that the time-dependent event lies within the interval (l_i, r_i) ($1 \leq i \leq n$). Assuming

A_1 Mechanism generating the random intervals (censorship), $(L, R]$, is independent from both, the time-to-event, T , and the marker, X , variables.

A_2 There exists a function, $\mathcal{S}_0(t)$, and a real number, β , such that, for a given value of the marker, x , then $\mathcal{S}(t|x) = \mathcal{S}_0(t)e^{x \cdot \beta}$.

Let $\mathcal{R}_t^{\text{C/D}}$ be the C/D ROC curve associated with the marker, X , at t satisfying

A_3 $\mathcal{R}_t^{\text{C/D}}$ has two continuous and bounded derivatives.

Then, if $\hat{\mathcal{R}}_t^{\text{C/D}}$ denotes the Mixed subject C/D ROC curve estimator based on the Finkelstein [1986] model, we have the uniform convergence

$$\sup_{p \in (0,1)} |\hat{\mathcal{R}}_t^{\text{C/D}}(p) - \mathcal{R}_t^{\text{C/D}}(p)| \xrightarrow{n} 0 \quad (\text{almost surely}).$$

Proof.

First we prove, under assumptions A_1 and A_2 , the uniform consistency of both the sensitivity and the specificity and then that these consistencies are enough for having the uniform consistency of the C/D ROC curve estimator based on those.

Let $\mathcal{I}_o(t) \subset \{1, \dots, n\}$ the set of indices containing those subjects in the sample whose real status at t are really observed ($\hat{p}_i(t) = 0$ or $\hat{p}_i(t) = 1$, $i \in \{1, \dots, n\}$) and let n_o be its cardinal. We have that, for each $x \in \mathbb{R}$,

$$\frac{1}{n} \sum_{i=1}^n I_{(x, \infty)}(x_i) \cdot \hat{p}_i(t) = \frac{n_o}{n} \cdot [1 - \hat{F}_t(x)] + \frac{1}{n} \sum_{i \notin \mathcal{I}_o(t)} I_{(x, \infty)}(x_i) \cdot \hat{p}_i(t),$$

where \hat{F}_t is the empirical cumulative distribution function (ECDF) estimator for the CDF of the marker values on positive subjects, F_t , based on those who we know they actually are positive at time t . Besides,

$$\frac{1}{n} \sum_{i \notin \mathcal{I}_o(t)} I_{(x, \infty)}(x_i) \cdot \hat{p}_i(t) = \int \int_x^\infty \hat{\mathcal{P}}(T \leq t \mid \mathbf{W}, X = x) \cdot [1 - I_{O_t}(\mathbf{W})] d\hat{H}_{\mathbf{W}} dx,$$

where \mathbf{W} stands for the random variable containing all the information available in the sample (excluding the marker values), being $H_{\mathbf{W}}$ its CDF and $\hat{H}_{\mathbf{W}}$ the corresponding ECDF and X depicts the random variable representing the marker. Denoting $\mathbf{Y} = (\mathbf{W}, X)$, $\hat{\mathcal{P}}(T > t \mid \mathbf{Y})$ stands for the predictive model on which the Mixed subject C/D ROC curve estimator is based and O_t the subset in which the information contained in \mathbf{Y} determines the real status at t . The Large Numbers Law guarantees that $n_o/n \rightarrow_n \pi_o(t)$ ($= \mathcal{P}(\mathbf{Y} \in O_t)$). On the other hand, from A_1 and A_2 , we directly have that

$$\mathcal{P}(X > x, T \leq t) = \pi_o(t) \cdot [1 - F_t(x)] + \int \int_x^\infty \mathcal{P}(T \leq t \mid \mathbf{W}, X = x) \cdot [1 - I_{O_t}(\mathbf{W})] dH_{\mathbf{W}} dx.$$

Finkelstein [1986] guarantees that, under assumptions A_1 and A_2 , for each $t > 0$, $|\hat{p}_i(t) - p_i(t)| \rightarrow_n 0$. The Glivenko-Cantely Lemma completes the argument for having that,

$$\sup_{x \in \mathbb{R}} \left| \frac{1}{n} \sum_{i=1}^n I_{(x, \infty)}(x_i) \cdot \hat{p}_i(t) - \mathcal{P}(X > x, T \leq t) \right| \rightarrow_n 0 \quad (\text{almost surely}),$$

and directly, we obtain the uniform convergence for the cumulative sensitivity, that is

$$\sup_{x \in \mathbb{R}} |\hat{S}e_t^{\mathbb{C}}(x) - Se_t^{\mathbb{C}}(x)| \rightarrow_n 0 \quad (\text{almost surely}). \quad (1)$$

Arguing similarly we get

$$\sup_{x \in \mathbb{R}} |\hat{S}p_t^{\mathbb{D}}(x) - Sp_t^{\mathbb{D}}(x)| \rightarrow_n 0 \quad (\text{almost surely}). \quad (2)$$

Now, we are proving that these convergences guarantee the C/D ROC curve consistency. From A_3 , for any sequence $\{p_n\}_{n \in \mathbb{N}} \in (0, 1)$, it holds the equality,

$$[\hat{\mathcal{R}}_t^{\text{C/D}}(p) - \mathcal{R}_t^{\text{C/D}}(p)] = [\hat{\mathcal{R}}_t^{\text{C/D}}(p) - \mathcal{R}_t^{\text{C/D}}(p_n)] + r_t(p) \cdot (p_n - p) + \mathcal{O}([p_n - p]^2).$$

where $r_t(\cdot)$ stands for the (bounded) first derivative of $\mathcal{R}_t^{\text{C/D}}$. Let be $x_n = [1 - \hat{S}p_t^{\text{D}}]^{-1}(p)$ and $p_n = 1 - Sp_t^{\text{D}}(x_n)$, then

$$\begin{aligned} [\hat{\mathcal{R}}_t^{\text{C/D}}(p) - \mathcal{R}_t^{\text{C/D}}(p)] &= [\hat{S}e_t^{\text{C}}(x_n) - Se_t^{\text{C}}(x_n)] \\ &\quad + r_t([1 - \hat{S}p_t^{\text{D}}](x_n)) \cdot [Sp_t^{\text{D}}(p) - \hat{S}p_t^{\text{D}}(p)] + [Sp_t^{\text{D}}(p) - \hat{S}p_t^{\text{D}}(p)]^2. \end{aligned} \tag{3}$$

Results in (1), (2) and (3) directly imply the strong uniform consistency.

Appendix B:

```
intCDroc <- function(data, t){
# Incoming parameters:
# data -> three-columns matrix: lower and upper edges of the observed
#       interval and marker values.
#   t -> numeric value representing the moment of time at which the C/D ROC
#       curve will be estimated.

# Split the incoming data matrix:
L <- as.vector(data[,1]) # Lower edge of the observable interval
R <- as.vector(data[,2]) # Upper edge of the observable interval
X <- as.vector(data[,3]) # Marker values

# Basic data checks:
if (sum(L > (ifelse(is.na(R), (max(L) + 1 ), R))) > 0){
  stop(message ("Non-valid data.
                Interval left side greater than right side"))}

# First auxiliary function implementing the Mixed subjects estimator
```

```

# for the computation of the probabilities of being positive (p_i)
# for the undefined individuals:
pundef <- function(i,M, mod, pt){

# Second auxiliary function for computing the value of the
# survival function estimate for a given point. If it belongs to a
# Turnbull Interval (TI), the estimate is obtained by interpolating
# the two consecutive survival function values before and after
# the TI. Othercase, the corresponding survival function value is taken.
pi <- function(p,M,S){
In <- which((M[,1]<= p) & (M[,2] >= p))
pi <- ifelse(length(In) == 0,
             ifelse(is.finite(max(which((M[,1]<= p))))),
             S[max(which((M[,1]<= p))],1),
             approxfun(c(M[In,1], M[In,2]),c(S[In-1], S[In]))(p))}

# Marker value
new_data <- data.frame(X = c(as.numeric(X[i])))

# Survival curve estimate corresponding the given marker value
rownames(new_data) <- c('grp1')
Ma <- getSCurves(mod,new_data)

# Mixed subjects estimator computation for p_i
d <- ifelse(is.na(M[i,2]),0,
            pi(M[i,2],Ma$Tbull_ints,Ma$S_curves$grp1))
p_i <-(pi(pt,Ma$Tbull_ints,Ma$S_curves$grp1) - d) /
      (pi(M[i,1],Ma$Tbull_ints,Ma$S_curves$grp1) - d)}

# Initialization of vectors
Pr <- NULL
auc <- NULL

```

```

# Computation of the proportional hazard model for interval-censored data
# (function ic_sp from icenReg package is used).
modsp <- ic_sp(Surv(L, R, type="interval2") ~ X ,
               B = c(0,1), model = "ph")

# Allocation of probabilities p_i
# Positive subjects; p_i <- 1
IL <- which (L >= t); Pr[IL] <- 1

# Negative subjects; p_i <- 0
IR <- which (R <= t); Pr[IR] <- 0

# Undefined subjects; p_i computed by the Mixed subjects estimator
# implemented in the pundef function.
IU <- which((L < t) & ( t < R | is.na(R)))
Pr[IU] <- sapply(IU, FUN = pundef, M = data, mod = modsp, pt = t)

# Computed probabilities are added to the original data matrix
data  <- cbind(data,Pr)

# False positive
FP    <- 1 -
      (sapply(X, function(i){sum(data[which(X <= i),4])}) /
       sum(data[,4]))

# True positive
TP    <- sapply(X, function(i){sum(1 - (data[which(X > i),4])}) /
      sum(1-data[,4])

# ROC curve
p     <- seq(0, 1, 1/50)

```

```

R      <- approxfun(c(0,FP,1),c(0,TP,1))(p)

# AUC (rectangles method)
auc   <- c(auc,(p[2] - p[1]) * sum(R))

# Function output: sROct object.
# sROct is a list with the following elements:
# th  <- vector of thresholds or cut-off points.
# FP  <- vector of false positive (complement to specificity).
# TP  <- vector of true positive (sensitivities).
# p   <- sequence of equidistant points between 0 and 1 for which C/D
#       ROC curve is plotted.
# ROC <- sequence of values of the ROC curve computed for each p.
# t   <- single value at which the C/D ROC curve is estimated.
# auc <- area under the C/D ROC curve.

ret = list( th = X, FP = FP,  TP = TP, p = p, ROC = R,  t = t, auc = auc)
class(ret) <- "sROct"
return(ret)}

```

References

- X.-H. Zhou, N. A. Obuchowski, and D. K. McClish. *Statistical Methods in Diagnostic Medicine*. Wiley Blackwell, New York, 2002.
- M. S. Pepe. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford Statistical Sciences Series, 2003.
- R. Fluss, D. Faraggi, and B. Reiser. Estimation of the Youden index and its associated cutoff point. *Biometrical Journal*, 20(4):458–472, 2005.
- P. Martínez-Camblor and J. C. Pardo-Fernández. Parametric estimates for the receiver

- operating characteristic curve generalization for non-monotone relationships. *Statistical Methods in Medical Research*, 28(7):2032–2048, 2019. doi: 10.1177/0962280217747009.
- P. Martínez-Cambor, N. Corral, C. Rey, J. Pascual, and E. Cernuda-Morollón. Receiver operating characteristic curve generalization for non-monotone relationships. *Statistical Methods in Medical Research*, 26(1):113–123, 2017.
- R. Etzioni, M. S. Pepe, G. Longton, C. Hu, and G. Goodman. Incorporating the Time Dimension in Receiver Operating Characteristic Curves: A Case Study of Prostate Cancer. *Medical Decision Making*, 19(3):242–251, 1999.
- M. S. Pepe, Yuying Zheng, Y.J., H. Ying, P. Chirag R., and W. C. Levy. Evaluating the ROC performance of markers for future events. *Lifetime Data Analysis*, 14(1):86–113, 2008.
- P. J. Heagerty, T. Lumley, and M. S. Pepe. Time-Dependent ROC Curves for Censored Survival Data and a Diagnostic Marker. *Biometrics*, 56(2):337–344, 2000.
- J. Li and S. Ma. Time-dependent ROC analysis under diverse censoring patterns. *Statistics in Medicine*, 30(11):1266–1277, 2011.
- H. Jacqmin-Gadda, P. Blanche, E. Chary, C. Touraine, and F. G. Dartigues. Receiver operating characteristic curve estimation for time-to-event with semicompeting risks and interval censoring. *Statistical Methods in Medical Research*, 25(6):2750–2766, 2016.
- P. J. Heagerty and Y. Zheng. Survival Model Predictive Accuracy and ROC Curves. *International Biometric Society*, 61(1):92–105, 2005.
- E. H. Slate and B. W. Turnbull. Statistical models for longitudinal biomarkers of disease onset. *Statistics in Medicine*, 4(19):617–637, 2000.
- P. Blanche, J. F. Dartigues, and H. Jacqmin-Gadda. Review and comparison of ROC curve estimators for a time-dependent outcome with marker-dependent censoring. *Biometrical Journal*, 55(5):687–704, 2013.
- L. Chambles and G. Diao. Estimation of time-dependent area under ROC curve for long-term risk prediction. *Statistics in Medicine*, 20(25):3474–3486, 2006.

- H. Uno, T. Cai, L. Tian, and L. J. Wei. Evaluating Prediction Rules for t-Year Survivors With Censored Regression Models. *Journal of the American Statistical Association*, 478(102):527–537, 2007.
- H. Hung and C. Chiang. Optimal composite markers for time-dependent receiver operating characteristic curves with censored survival data. *Scandinavian Journal of Statistics*, 20(37):664–679, 2010.
- P. Wolf, G. Schmidt, and K. Ulm. The use of ROC for defining the validity of the prognostic index in censored data. *Statistics & Probability Letters*, 20(81):783–791, 2011.
- P. Martínez-Camblor and J. C. Pardo-Fernández. Smooth time-dependent receiver operating characteristic curve estimators. *Statistical Methods in Medical Research*, 27(3):651–674, 2018.
- P. Martínez-Camblor, G. F. Bayón, and S. Pérez-Fernández. Cumulative/dynamic ROC curve estimation. *Journal of Statistical Computation and Simulation*, 86(17):3582–3594, 2016.
- Y. Wu, X. Wang, J. Lin, J. Beilin, and K. Owzar. Predictive Accuracy of Markers or Risk Scores for Interval Censored Survival Data. *arXiv:1806.01760v1 [stat.ME]*, 2018.
- L. Li, T. Greene, and B. Hu. A simple method to estimate the time-dependent receiver operating characteristic curve and the area under the curve with right censored data. *Statistical Methods in Medical Research*, 27(8):2264–2278, 2018.
- D. M. Finkelstein. A Proportional Hazards Model for Interval-Censored Failure Time Data. *Biometrics*, 42(4):845–854, 1986.
- B. W. Turnbull. The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 38(3):290–295, 1976.
- X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J. C. Sánchez, and M. Müller. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12(3):77, 2011.

- C. Anderson-Bergman. icenReg: Regression Models for Interval Censored Data in R. *Journal of Statistical Software*, 81(12):1–23, 2017. doi: 10.18637/jss.v081.i12.
- G. Gómez, M. L. Calle, R. Oller, and K. Langohr. Tutorial of methods for interval-censored data and their implementation in R. *Statistical Modelling*, 9(4):259–297, 2009.
- D. Lavanchy. The global burden of hepatitis C. *Liver International*, 29 Suppl 1:74–81, 2009.
- J. H. Hoofnagle. Course and outcome of hepatitis C. *Hepatology*, 36:21–29, 2002.
- K. Matsuzaki and M. and Yoshida K. Murata. Chronic inflammation associated with hepatitis C virus infection perturbs hepatic transforming growth factor beta signaling, promoting cirrhosis and hepatocellular carcinoma. *Hepatology*, 46:48–57, 2007.