# The International Journal of Biostatistics

Volume 8, Issue 1

2012

Article 1

# Cut-Off Estimation and Medical Decision Making Based on a Continuous Prognostic Factor: The Prediction of Kidney Graft Failure

Yohann Foucher, Nantes University
Magali Giral, INSERM U643
Jean-Paul Soulillou, INSERM U643
Jean-Pierre Daurès, University Institut of Clinical Research

#### **Recommended Citation:**

Foucher, Yohann; Giral, Magali; Soulillou, Jean-Paul; and Daurès, Jean-Pierre (2012) "Cut-Off Estimation and Medical Decision Making Based on a Continuous Prognostic Factor: The Prediction of Kidney Graft Failure," *The International Journal of Biostatistics*: Vol. 8: Iss. 1, Article 1.

# Cut-Off Estimation and Medical Decision Making Based on a Continuous Prognostic Factor: The Prediction of Kidney Graft Failure

Yohann Foucher, Magali Giral, Jean-Paul Soulillou, and Jean-Pierre Daurès

#### **Abstract**

The determination of a cut-off value for a continuous prognostic test is an important problem, which is statistically challenging and practically important for risk assessment. We propose in this paper a method to estimate the optimal cut-off from this type of longitudinal data with censored failure times. The principle is to combine the prognostic error rates of false positives and false negatives with a cost function, which has the advantages to be statistically convenient and to be directly associated with the decision-making. Simulations were performed and the results demonstrate the interest of our approach compared to a reference method. The method is also illustrated by predicting the long-term survival of kidney transplant recipients from the 1-year creatinine clearance.

KEYWORDS: cut-off estimation, censored data, prognostic test, decision-making, cost function

**Author Notes:** Yohann Foucher is affiliated with the Institut of Transplantation and Research in Transplantation - Department of Biostatistics EA 4275 - Nantes University.

Magali Giral is affiliated with the Institut of Transplantation and Research in Transplantation - INSERM U643.

Jean-Paul Soulillou is affiliated with the Institut of Transplantation and Research in Transplantation - INSERM U643.

Jean-Pierre Daurès is affiliated with the University Institut of Clinical Research - Unit of Biostatistics.

We thank the Roche Laboratory for its support of the DIVAT data bank and Dr J. Ashton-Chess for editing the manuscript. This work is a part of the French Transplantation Research Network (RTRS) supported by the CENTAURE foundation.

### 1 Introduction

Cut-off determination is a frequent issue for statisticians. In clinical or epidemiological studies, the effect of continuous risk factors are often analyzed respecting a dose-effect relationship [Sebaugh et al., 1991]. For example in kidney transplantation, numerous studies demonstrated that creatinine clearance is highly correlated with long-term graft survival [Nicol et al., 1993, Giral et al., 1996, Hariharan et al., 2002]. However, it is difficult for clinicians to apply these results in practice as no cut-off is estimated for a decision. Moreover, from a statistical point of view, it may be important to categorize continuous covariates when a dose-effect assumption does not hold.

In order to determine such cut-offs in survival analyses with censored followup, a widely used method is to define a grid and to retain the cut-off associated with the highest difference between survival curves. However, with such a procedure, investigators are confronted with the problem of multiple testing. Therefore, authors such as LeBlanc and Crowley [1993] have considered tree methods. These procedures split data by maximizing the difference in survival between groups, which is commonly measured by the log-rank test. Contal and O'Quigley [1999] have also proposed the maximization of a statistic, which bears the advantage of avoiding cut-offs located near the extremes. The most recent publication on this topic is based on the generalized maximally selected statistics [Hothorn and Zeileis, 2008]. The latter authors have proposed an algorithm for a unified treatment of different kinds of maximally selected statistics enabling a large number of cut-offs. The procedure leads to a maximally selected chi-square test, as published by Miller and Siegmund [1982]. The common characteristic of these methods is their independence to their application. However, risk assessment in medical practice has to take into account the consequences of the clinician decision.

In this paper, we propose an alternative method for cut-off estimation based on a decision-making framework. Depending on the application, this method reflects the impact of prognostic errors (monetary cost, medical gravity, social consequences, etc.) and takes into account the desired time of the prognostic.

In Section 2, we describe the adaptation of time-dependent ROC curves, initially defined by Heagerty et al. [2000], for the estimation of cut-offs. We define in detail the ROC analysis because the evaluation of the prognostic capacity of a marker is of prime importance before the determination of a particular cut-off. Two non-parametric methods are proposed, based on the Kaplan-Meier estimator and on the Akritas nearest neighborhood estimator. In Section 3, we propose comparing the recent method of Hothorn and Zeileis [2008] with the new approaches by simulations. Section 4 applies the methods

to the analysis of kidney transplant recipients. The clinical objectives are the evaluation of the prognostic capacity of the 1-year creatinine clearance (CrCl) and the definition of the optimal cut-off to discriminate two groups according to their risk of failure. Finally, Section 5 discusses the new method and its benefits and limitations.

## 2 Methods

#### 2.1 Framework

Using the counting process notation, let D(t) = 1 if the failure occurred before time t (i.e., T < t) and D(t) = 0 otherwise (i.e., T > t). All patients are free of failure at the beginning of the study (D(0) = 0). We consider X, measured at t=0, as a prognostic marker of the failure time T. By convention, suppose that high values of X are associated with a high risk of failure. The prognostic test is defined as positive (patient at risk of failure), if the prognostic marker is higher than a cut-off c. The methodology associated with this type of prognostic analysis has been recently developed by Heagerty et al. [2000, 2005. The sensitivity is thus the probability of a positive prognostic test among patients with failure before time t, i.e. P(X > c|D(t) = 1). The specificity is the probability of a negative prognostic test among patients free of failure before time t, i.e.  $P(X \le c|D(t) = 0)$ . The ROC (Receiver Operating Characteristic) curve of a prognostic at time t represents the sensitivity in function of one minus the specificity for the different cut-offs c. The ROC curve is monotone non-decreasing for each t. The accuracy of the marker to predict the failure is measured by the area under the curve (AUC), independently of the cut-off. In order to find the optimal cut-off, we define a cost function, C(c|t), which represents the total cost associated with the prognostic test based on the cut-off c at time t. One can distinguish the number of false positives (patients with a positive test but free of failure at time t) and the number of false negatives (patients with a negative test but with a failure before time t), respectively  $n_{(fp)}(c|t)$  and  $n_{(fn)}(c|t)$ . As a result, the cost function represents the sum of these errors weighted by their respective costs,  $C_{(fp)}$  and  $C_{(fn)}$ . Thus, we have:

$$C(c|t) \propto k \times n_{(fp)}(c|t) + n_{(fn)}(c|t), \tag{1}$$

2

where  $\propto$  means "proportional to" and where  $k=C_{(fp)}/C_{(fn)}$  is the relative importance of a false positive according to a false positive. Depending on the application, this function C(c|t) can be determined such that it reflects the monetary impact, the medical gravity, the social impact, etc. The optimal cut-off is calculated to minimize this cost function for a given time of prognosis t.

### 2.2 Estimation using the Kaplan-Meier estimator

In the framework above, the sensitivity of the prognostic test with cut-off c at time t represents the probability of having a positive test  $\{X > c\}$ , given that a failure occurs before time t. Based on the work of Heagerty et al. [2000], the non-parametric estimation of this probability,  $se_{KM}(c|t)$ , is:

$$se_{KM}(c|t) = \{1 - S_{KM}(t|X > c)\}\{1 - \bar{G}_X(c)\}/\{1 - S_{KM}(t)\},$$
 (2)

where  $S_{KM}(t|X>c)$  is the Kaplan-Meier estimator of the survival probability at time t conditional on  $\{X>c\}$  and  $\bar{G}_X(c)$  is the empirical distribution function. Respectively, the estimator of the specificity of a prognostic test with cut-off c at time t,  $sp_{KM}(c|t)$ , is:

$$sp_{KM}(c|t) = S_{KM}(t|X \le c)\bar{G}_X(c)/S_{KM}(t) \tag{3}$$

The ROC curve of a prognostic at time t, is the sensitivity  $se_{KM}(c|t)$  plotted in function of one minus the specificity  $sp_{KM}(c|t)$  for all the possible cut-off c. In order to find the optimal cut-off,  $\tilde{c}$ , the cost function (1) can be developed as follow:

$$C(c|t) \propto kP(X > c, D(t) = 0) + P(X \le c, D(t) = 1)$$
  
  $\propto kP(T > t|X > c)P(X > c) + P(T \le t|X \le c)P(X \le c),$ 

and can be non-parametrically estimated:

$$C_{KM}(c|t) \propto kS_{KM}(t|X>c)\{1-\bar{G}_X(c)\} + \{1-S_{KM}(t|X\leq c)\}\bar{G}_X(c)$$
 (4)

# 2.3 Estimation using the Akritas estimator

Heagerty et al. [2000] also proposed to estimate P(X > c, D(t) = 0) using the nearest neighborhood estimator initially proposed by Akritas [1994]. Respecting the same notations, let  $S_{\lambda_n}(c,t)$  denote the estimation of this bivariate survival probability, where  $2\lambda_n \in (0,1)$  represents the percentage of observations that is included in each neighborhood. This estimator ensures a monotone ROC curve in contrast to the Kaplan-Meier approach. Moreover, the Kaplan-Meier estimator will be biased if censoring is dependent on marker, whereas the Akritas estimator will be robust in this situation. The sensitivity and specificity become:

The International Journal of Biostatistics, Vol. 8 [2012], Iss. 1, Art. 1

$$se_{\lambda_n}(c|t) = \{(1 - \bar{G}_X(c)) - S_{\lambda_n}(c,t)\}/\{1 - S_{\lambda_n}(-\infty,t)\}$$
 (5)

$$sp_{\lambda_n}(c|t) = 1 - S_{\lambda_n}(c,t)/S_{\lambda_n}(-\infty,t)$$
 (6)

The ROC curve of a prognostic at time t is similarly obtained using the probabilities (5) and (6) instead of (2) and (3). The cost function can be estimated by:

$$C_{\lambda_n}(c|t) \propto (k+1)S_{\lambda_n}(c,t) + \bar{G}_X(c)$$
 (7)

### 2.4 Computation details

We chose the trapezoid method to calculate the area under the curve. In order to calculate the 95% confidence intervals  $(CI_{95\%})$  of the optimal cut-offs and of the areas under the curves, 1999 bootstrap replications were performed and percentile intervals were calculated [Efron, 1987].

## 3 Results from simulations

Both approaches were compared to the Hothorn and Zeiles method [2008], which identifies a global cutoff over all failure time and is based on the maximization of the difference between the survival curves (generalized maximally selected statistics). The false positive and negative errors are not taken into account. The Hothorn and Zeiles method is thus completely independent of the application and the decision consequences. Therefore, the present comparisons are only relevant if no assumption is made about the prognostic time and the consequences. We arbitrarily chose k = 1 (no preference regarding the minimization of the false positive or the false negative errors). We also arbitrarily chose a prognostic time equal to half the maximum observed follow-up time.

We simulated artificial samples for different sample sizes (N=25, 50, 100 and 200). Let us suppose a variable Z simulated assuming a standard normal. This variable was used in the 4 following scenarios in order to define the value of X:

1. The times-to-event are simulated according to a proportional hazard model with a Weibull distribution (scale and shape parameters respectively equal to 1.5 and 0.5). z is the observation of a random variable Z which is a transformation of X. We assume that the regression coefficient  $\beta$  equals 0.4 (relative risk equals 1.5). A single cut-off is fixed

4

at 0.5 with Z = 1 if X > 0.5 and with Z = 0 otherwise. The times of censoring are uniformly distributed between 0 and 15.

- 2. The times of events are simulated similarly. However, we fix a cut-off in 0.5 with X=Z if Z>0.5 and with X=0 otherwise. In contrast to the scenario #1, the hazard function does not jump at 0.5. The function is constant before 0.5 and proportionately increases with the values above 0.5.
- 3. If  $Z \leq 0.5$  the times of events are simulated according to a Weibull law (scale and shape parameters respectively equal to 1.5 and 0.5), but if Z > 0.5 the Exponential distribution is used (scale and shape parameters respectively equal to 1.5 and 1.0). Thus, a single cut-off exists at 0.5 but the PH assumption does not hold.
- 4. The scenario is equivalent to (1) but with X = Z. There is no cut-off.

1000 simulations were performed per scenario and per sample size. The results are presented in Table (1). Globally, the results were very similar regardless of the method. Below a sample size of 50 individuals, no method was reliable, but the proposed methodology (Kaplan-Meier or Akritas) offered estimations closer to the true cut-off in comparison to the Hothorn and Zeileis method. Regardless of the scenario, the variability of the estimations using cost function seemed to be lower than those obtained by the Hothorn and Zeileis method. If no cut-off exists (Scenario #4), the proposed method estimated cuts-offs close to 0, which separated the sample into two balanced groups. In this scenario, the Hothorn and Zeileis method offered lower estimations.

# 4 Kidney transplant survival data

## 4.1 Data description

The data were extracted from the DIVAT data bank from Nantes Hospital (France), which is a prospective cohort of kidney transplant recipients. Biological and clinical data have been recorded since 1990. Specialized clinical research assistants who were independent to the medical team, computerized the pre- and post-transplant parameters of each patient transplanted in the center. Recorded data are submitted to an annual medical cross-audit with a level of error below 1%.

In this paper, we consider a subpopulation of 839 patients more than 18 years of age and who received a kidney transplant between January 1996 and

Size	H. & Z.	K. & M.	Akritas
		Scenario #1	
N = 25	-0.06 (1.07)	0.10 (0.97)	0.11 (0.98)
N = 50	0.13 (1.09)	0.28 (0.83)	0.27(0.84)
N = 100	0.30 (0.82)	0.36 (0.73)	0.34(0.69)
N = 200	0.41 (0.51)	0.38 (0.46)	0.38 (0.43)
		Scenario #2	
N = 25	0.02 (1.20)	0.21 (0.98)	0.18 (1.08)
N = 50	0.25(1.08)	0.36 (0.90)	0.33(0.89)
N = 100	0.49(0.81)	0.41 (0.73)	0.41(0.69)
N = 200	0.57 (0.50)	0.49(0.50)	0.49(0.50)
		Scenario #3	
N = 25	0.24 (1.05)	0.16 (0.85)	0.17 (0.80)
N = 50	0.36 (1.01)	0.32(0.63)	0.32(0.60)
N = 100	0.44(0.77)	0.40 (0.44)	0.40 (0.40)
N = 200	0.49(0.37)	0.43 (0.26)	0.43 (0.24)
		Scenario #4	
N = 25	-0.28 (1.09)	0.05 (0.83)	-0.04 (0.81)
N = 50	-0.35 (1.06)	0.00(0.76)	0.01 (0.74)
N = 100	-0.32 (0.91)	-0.01 (0.70)	-0.03 (0.67)
N = 200	-0.31 (0.80)	0.01 (0.61)	0.02 (0.60)

Table 1: Estimations of the cut-off values according to the scenarios and the sample sizes. Based on 1000 simulations, the medians (and the corresponding inter-quartile intervals) are reported.

September 2006. Death or a return to dialysis mean that the graft failed. The value of the creatinine clearance (CrCl) one year after the transplant is the marker of interest to predict long-term graft survival. A low CrCl value is associated with a higher risk of graft failure. Note that there is no reason to justify that a cut-off exists with a discontinuity of the risk at this value. However, clinicians have to take decisions to classify the patients according to their risk of failure. Usually, clinicians consider a CrCl above  $40 \, ml/min$  as an indication of a poor prognosis. A cut-off of 40 ml/min is obtained as half of the lower CrCl limit for a healthy person (80 ml/min). When one kidney is transplanted to a patient, one expects a CrCl of half the normal values. To our knowledge, no quantitative study has been performed to justify this threshold according to the risk of failure. Based on the Kaplan-Meier analysis of graft survival, Hariharan et al. have shown that a creatinine level of more than 1.5 mg/dL is associated with a poor graft outcome, compared to levels below that value [Hariharan et al., 2002]. The focus of interest is whether this cut-off is optimal for discriminating two groups of patients according to their risk of graft failure.

Our objective was to determine the optimal cut-off of the CrCl value collected 1 year after the transplantation for predicting graft survival. The origin of the study (t=0) is thus at 1 year after the transplantation and concerns only patients with a functional kidney at 1 year. Return to dialysis, death and censored patient within the first year of transplantation were not included in the analysis. The survival time of interest is thus the time between the first anniversary of the transplantation and the graft failure or the death of the patient. In the following developments, we will principally choose a prognostic time of up to 4 or 8 years after the CrCl measurement (t=4,8). Different formulae are available to calculate the CrCl. We used the Modification of Diet in Renal Disease version [Levey et al., 1999]. The mentioned data are provided for reanalysis and verification (http://www.divat.fr/).

#### 4.1.1 Study of the prognostic accuracy

Figure 1 shows the ROC curves based on the Kaplan-Meier or on the Akritas estimators. It illustrates the ability of the CrCl to predict a failure up to 4 and up to 8 years after the first anniversary of the transplantation. In agreement with Heagerty et al. [2000], both estimators gave similar results. Using the Akritas version, the AUC at 4 years was 0.79 ( $CI_{95\%} = [0.70, 0.85]$ ) versus 0.73

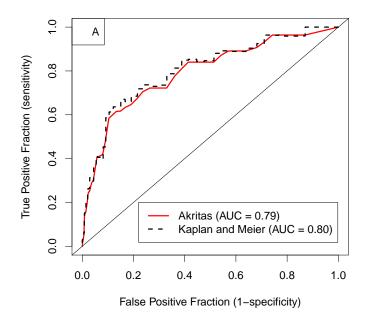
 $(CI_{95\%} = [0.69, 0.87])$  at 8 years. Using the Kaplan-Meier estimator, the AUC at 4 years was 0.80  $(CI_{95\%} = [0.69, 0.87])$  versus 0.70  $(CI_{95\%} = [0.50, 0.82])$  at 8 years. Regardless of the estimators and as expected, the CrCl was a better prognostic variable for a short-term prognostic.

#### 4.1.2 Determination of an optimal cut-off

If the objective of the cut-off is to discriminate all the events with a minimum of errors, then k should be equal to 1 (the cost of a false negative is equivalent to the cost of a false positive). However, in such a case, the cut-off can appear to be disconnected from the real medical issue. In our application, the medical consequences of a false negative are much more serious than the consequences of a false positive. The relative weight k is assumed at 1/9. We used the approach of Vickers and Elkin [2006] to determine this ratio. Consider that the follow-up of the at-risk patients is more frequent (every 3 months) in comparison with the follow-up of the risk-free patients (every year). If the probability of failure is close to 1, all clinicians will decide on an intensive follow-up. If the probability of failure is close to 0, all clinicians will decide on a less intensive follow-up. After discussions, clinicians defined the disease probability for which the decision is unsure at 10%. In other words, if the probability of a graft failure is below 10%, they accept to increase the length of the intervals between two visits. Above 10%, they accept to decrease this length. Their decision appears unsure at about 10% and Vickers and Elkin [2006] demonstrated that the relative harms of a false positive and a false negative is thus equal to 10/(100-10).

Figure 2 represents the optimal cut-offs for k=1/9, for both estimators and for all the prognostic times. Consider a prognostic up to 4 years after the CrCl measurement, which corresponds to the time origin, 1 year after the transplantation. The optimal cut-off equals  $31.4 \, ml/min \, (CI_{95\%} = [26.7, 33.9]$ , Akritas estimator). For a prognostic up to 8 years after the CrCl measurement, the optimal cut-off also equals  $31.4 \, ml/min \, (CI_{95\%} = [27.7, 34.1]$ , Akritas estimator). One can see that these estimations are different from the usual cut-off of 40 ml/min. However, the cut-offs do not vary according to the prognostic time. The results based on the Kaplan-Meier estimator are very similar.

We did not compare our results with the Hothorn and Zeileis method, because we chose the specific value of k and the specific time of prognosis. The comparison would not be relevant.



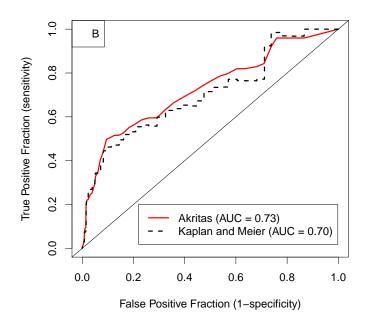


Figure 1: The ROC curves estimated for evaluating the capacity of the CrCl to predict a graft failure or the death of the transplanted recipient up to 4 years (A) and 8 years (B)

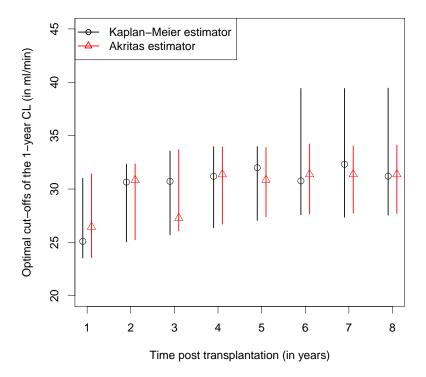


Figure 2: Optimal cut-offs depending on both estimators (Kaplan-Meier and Akritas) for k=1/9. The cut-off estimations are associated with the corresponding 95% confidence interval.

10

## 5 Discussion

In medicine, it is useful to determine a decision cut-off of prognostic markers. We have proposed a method in order to define such thresholds from time-dependent data where the failure may be censored.

We first proposed to validate the methodology using simulated data and comparing the results with the recent method of Hothorn and Zeileis [2008]. This approach is today a reference in traditional statistical analysis: the principal objective is the study of the correlation between risk factor and time-to-event and the estimation of cut-off associated with the lower p-value (null hypothesis that the hazard ratio equals 1). With the present approach, we distinguish the risk factor analysis from the prognostic analysis where the cut-off estimation needs to be adapted to the medical question. Therefore, for a relevant comparison of the methods, no assumption was made about the consequences of the prognostic decision. The objective of these simulations was to validate our methodology in this context. Regarding the simulation results, this objective was achieved. However, this does not mean that the proposed methodology is better. As explained, the methodologies should not be applied in the same context.

In accordance with Gail and Pfeiffer [2005], we believe that existing methods are irrespective of the intended application. Our proposed approach is more adequate in the context of medical decision making. Two possible modulations can be considered in order to obtain a cut-off close to the expectations of clinicians. Firstly, we can calculate the predictive accuracy and the cut-off value according to the required time of the prognosis. This distinction is of prime importance since a marker could be informative for early events, but not so useful for a long-term prediction. Our application illustrated this statement. Moreover, the fixing of the prognostic time is very important since the gravity of error may be different if it appears just after the decision or if it appears a long time after. From a statistical point of view, dichotomization of the survival time of the proposed method may lead to a loss of information. However, regarding the practice of the medical decision, determining the prognostic time is important.

Secondly, the concrete consequences of decision errors based on the cut-off can be taken into account. For example in transplantation, clinicians prefer to wrongly predict a future failure as opposed to wrongly predict the survival of the graft. This second error is only associated with a more intensive but useless follow-up. The priority is thus to minimize the number of false negatives (k < 1). The weights of both errors should be defined according to experts. The methodology proposed by Vickers and Elkin [2006] can be used to help

this definition. Even if no idea about the weights results from the discussion, the non-informative choice is always possible (k = 1).

The simulations demonstrated the good capacity of the methods to estimate the existing cut-off. One can ask the relevance of these estimations when no cut-off exist (scenario #5). However, in the medical practice, decisions have to be made even if no cut-off exists. This is the case in our application: the risk of graft failure decreases continuously with an increase in CrCl. Nevertheless, clinicians have to make decisions daily based on this marker.

A limitation of the proposed methodology is that no adjustment is possible to take into account confounding factors. The cut-off value can vary according to other determinants. However, based on the proposed nonparametric methodology, a solution is to perform a stratified analysis. To avoid the traditional limitations associated with stratified analysis, it may be interesting to develop a multivariate approach, semiparametric method can be considered.

Finally, it may also be of interest to generalize the method to multiple variables and to consider the marker as time-dependent. We are in the process of working on this type of extension, in particular based the recent paper by Zheng and Heagerty [2007].

## References

- MG Akritas. Nearest neighbor estimation of a bivariate distribution under random censoring. *Annals of Statistics*, 22:1299–1327, 1994.
- C Contal and J O'Quigley. An application of changepoint methods in studying the effect of age on survival in breast cancer. *Computational Statistics & Data Analysis*, 30:253–70, 1999.
- B Efron. Better bootstrap confidence intervals (with discussion). *Journal of the American Statistical Association*, 82:171–200, 1987.
- M Gail and R Pfeiffer. On criteria for evaluating models of absolute risk. *Biostatistics*, 6:227–39, 2005.
- M. Giral, C. Taddei, JM. Nguyen, J. Dantal, M. Hourmant, D. Cantarovich, G. Blancho, D. Ancelet, and JP. Soulillou. Single-center analysis of 468 first cadaveric kidney allografts with a uniform ATG-CsA sequential therapy. *Clinical Transplantation*, pages 257–64, 1996.

DOI: 10.2202/1557-4679.1215

12

- S Hariharan, MA McBride, WS Cherikhn, CB Tolleris, BA Bresnahan, and CP Johnson. Post-transplant renal function in the first year predicts long-term kidney transplant survival. *Kidney International*, 62:311–18, 2002.
- PJ Heagerty and Y Zheng. Survival Model Predictive Accuracy and ROC Curves. *Biometrics*, 64:92–105, 2005.
- PJ Heagerty, T Lumley, and SP Pepe. Time-Dependent ROC Curves for Censored Survival Data and a Diagnostic Marker. *Biometrics*, 56:337–44, 2000.
- T Hothorn and A Zeileis. Generalized Maximally Selected Statistics. *Biometrics*, 2008.
- M LeBlanc and J Crowley. Survival trees by goodness of split. *Journal of the American Statistical Association*, 88:457–67, 1993.
- AS Levey, JP Bosch, JB Lewis, T Greene, N Rogers, and D Roth. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Annals of Internal Medicine*, 130:461–70, 1999.
- R. Miller and D. Siegmund. Maximally Selected Chi Square Statistics. *Biometrics*, 38:1011–1016, 1982.
- D. Nicol, AS. MacDonald, J. Lawen, and P. Belitsky. Early prediction of renal allograft loss beyond one year. *Transplantation International*, pages 153–7, 1993.
- J. Sebaugh, J. Wilson, M. Tucker, and W. Adams. A Study of the Shape of Dose-Response Curves for Acute Lethality at Low Response: A Megadaphnia Study. *Risk Analysis*, 11:633–640, 1991.
- AJ Vickers and EB Elkin. Decision curve analysis: a novel method for evaluating prediction models. *Medical Decision Making*, 26:565–74, 2006.
- Y Zheng and PJ Heagerty. Prospective Accuracy for Longitudinal Markers. *Biometrics*, 63:332–41, 2007.