Article is available online at http://www.webio.hu/por/2005/11/1/0050

METHODS

ROC Analysis as an Additional Method to Characterize Time to Event Data

Horst J KOCH, Peter HAU

Neurologische Universitätsklinik Regensburg, Regensburg, Germany

The receiver operating characteristic (ROC) is predominantly used to assess the discriminant power of diagnostic tests. The present paper proposes this method as an additional alternative to compare survival data of in oncology or related fields. Survival data of brain tumors were analysed with conventional Kaplan-Meier method and ROC. The Area und the Curve (ROC-curve) gives additional illustrative information to distinguish between two therapeutic approaches and low or high grade brain tumors. (Pathology Oncology Research Vol 11, No 1, 50–52)

Key words: ROC, Kaplan-Meier, brain tumor

Introduction

The ROC (receiver operating characteristic) analysis method is widely used to compare two competing diagnostic tests. It also allows to determine optimal cut-off values of those tests (for mathematical basics see:[1,2]). In general, the procedure can be used, if a continuous variable (predictor, independent) may allow do distinguish between healthy or disease (binary variable). In some way, the procedure improves repeated crosstabs for different cut-off values of a selected parameter. This brief communication provides evidence that the method may be suited for additional application in survival analysis.

Theoretical Background

Survival data or more general time to event data are normally analysed using the Kaplan-Meier-Method[4] and can be illustrated using survivorship versus time plots for one or more variables. Characteristics such as mean survival time or corresponding quantiles can be derived from the the step function. Stratification variables („grouping“) may be introduced and the log rank test being the most popular test to look for significant differences.[3]

Received: Oct 20, 2004; *accepted:* Dec 1, 2004

Correspondence: Dr. med. Dr. rer. nat. Horst J. KOCH MFPM DCPSA, Neurologische Universitätsklinik am Bezirksklinikum Regensburg, Universitätsstrasse 84, 93053 Regensburg, Fax: 0941-941-3155. Email: horst.koch@medbo.de

In order to visualize the survival data with regard to theoretical (flip of the coin) curve, the ROC curve may be suited. The ROC analysis was primarily developed for signal detection of Radars. Nowadays, a common application is to judge between two diagnostic tests and to find an optimal cut-off value. Apart from the a list of specificities and sensitivities for subsequent cut-off values, an advantage of the method is the presentation is the ROC curve in a 1 *1 unit (or 100 *100 %) diagram. In this diagram sensitivity as plotted as function of 1-specifity. In more general terms one may also speak of „Hit rate“ versus „False Alarm Rate“. The length of abcissae and ordinates are normalized to one (1-specifity, sensitivity) and therefore the maximum area of the plot is per definitionem 1. The theoretical random curve starts at P(0,0) and ends at P(1,1) and the diagonal line separates the area in two halves having 0.5 units each. The Area under the ROC curve therefore lies between 0.5 and 1. The higher the AUC, the better the diagnostic test. In some way, the AUC could be interpreted as a probability, i.e. a selected patient, who has the disease, will have a higher (more pathological) test result compared to a control subject. An AUC of 0.7 means that a patient will have more pathological test results than 70% of healthy subjects. Although the AUC may be subject of test statistics and p-values may be calculated, this is not the purpose of this brief communication. One advantage of ROC curves that the AUC is strictly normalized to one and therefore allows to compare time to event curves of different population sizes or survival or event times. The suitability for separation of two states by a method or the difference of survival for a factor is therefore projected onto this 1*1 diagram.

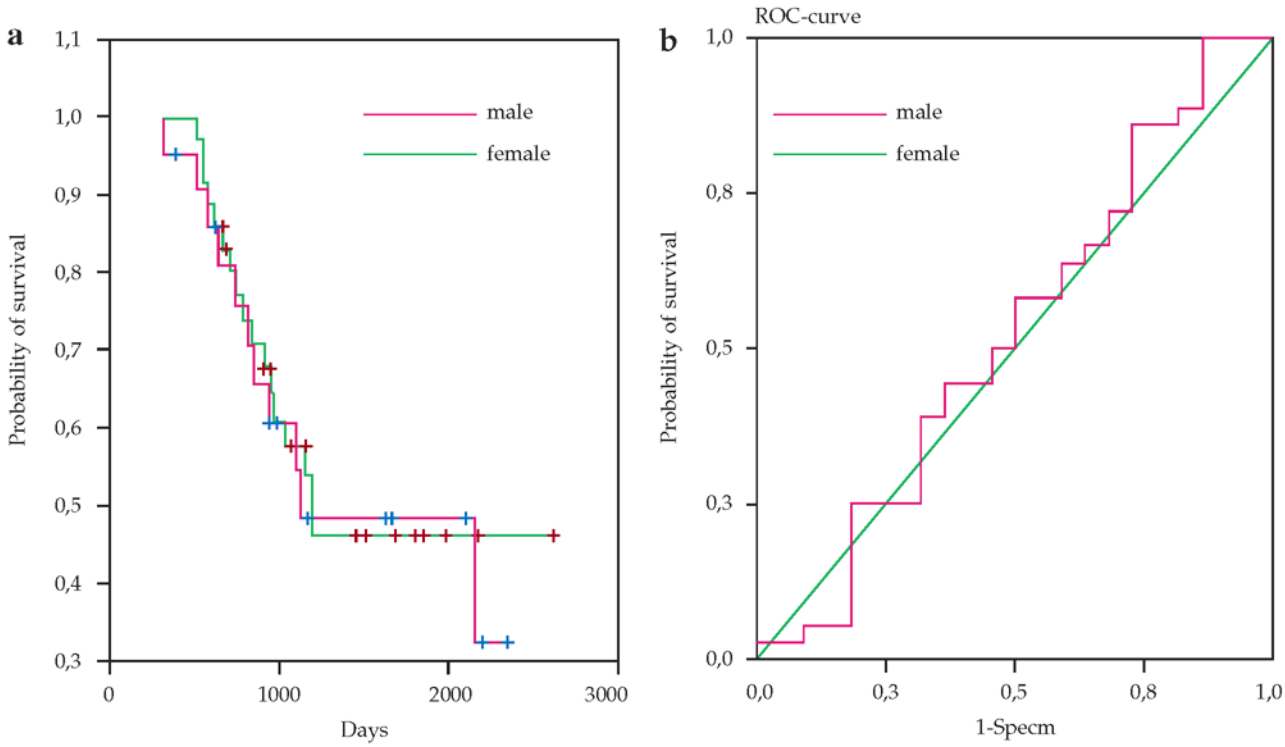
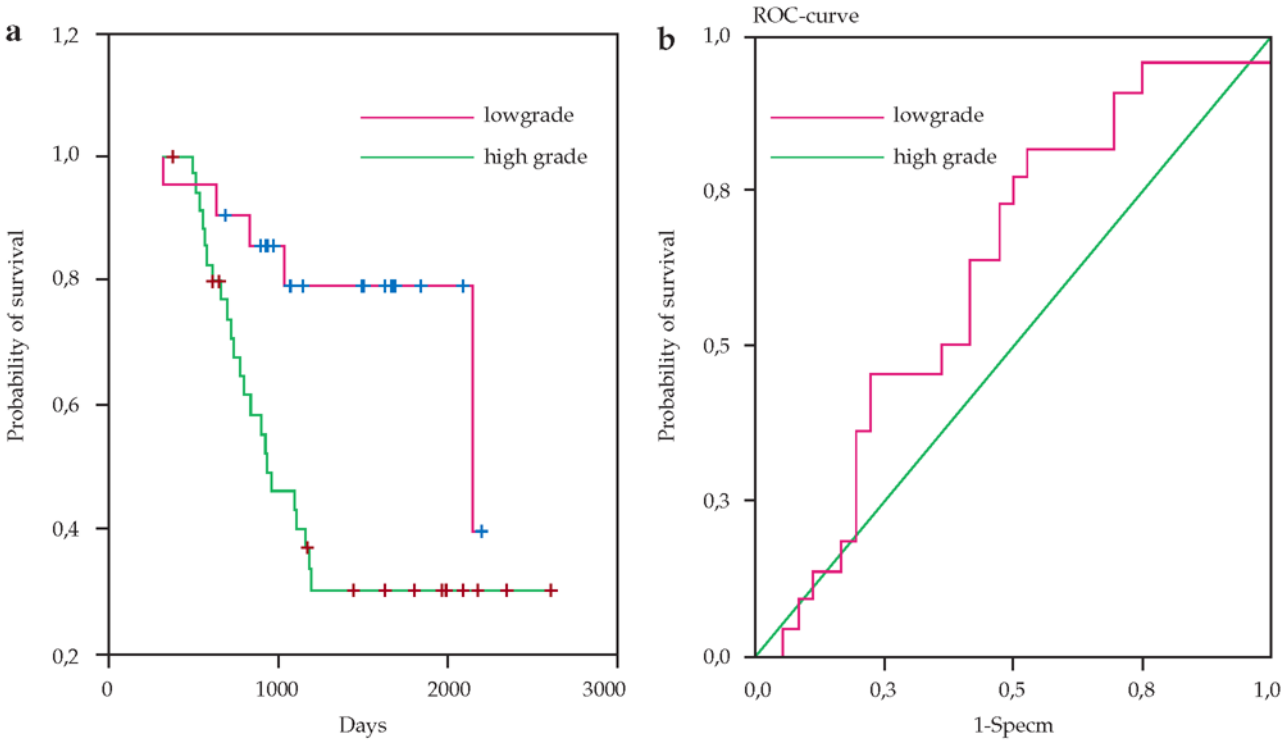


Figure 1. *Kaplan-Meier survival curves of high grade and low grade malignomas (mean survival times: 1659 (95%CI: 1343-1975) and 1505 (95% CI: 1157-1853) days). The factor „gender" does not have a significant impact on the prognosis (p>0.05, log rank test, nonparametric). The AUC of the corresponding ROC plot is 0.523 (95%CI: 0.363-0.683), which confirms the low „diagnostic" value of the factor „gender". (Remark: all data were entered until death or censored „+")*



Figures 2. *Kaplan-Meier survival curves of malignomas of both men and women (mean survival times: 1875 (95% CI: 1611-2138) and 1360 (95%CI; 1070-1650) days). The factor „high grade" does significantly separate both groups (p<0.05, log rank test, nonparametric). The corresponding ROC curves yield an AUC of 0.620 (95%CI: 0.473-0.767). Although the log rank test showed a significant result, the AUC does not indicate a substantial better overall prognosis of lower grade tumors.*

Materials and Methods

In order to illustrate the association between conventional survival analysis and the ROC method, survival data of malignant brain tumors [N=58, Age: 43.8 years (SD 10.4), 36 men, 22 women] different grades and treatments) were evaluated. ROC curves were calculated using a commerically available software package (Systat Version 11, Systat Software GmbH, Erkrath, Germany). The first theoretical test question was whether there is a gender difference of survival.

Results

The respective Kaplan-Meier curves are given in *Figure 1a*. The corresponding log rank test shows that there is no substantial difference between both sexes. This slight difference is also confirmed by the ROC curves (*Figure 1b*). The AUC is only 0.523, which corresponds to the negleg-ible diagnostic value of the parameter „gender". The second test question looks investigates the impact of the grade of malignancy on the the outcome. High grade (IV) tumors were compared with low grade tumors (III, II). Now, there is a significant difference (log rank rest) and the Kaplan Meier curves are quite separated (*Figure 2a*). The ROC curves also indicate a difference between low and high grade tumors (*Figure 2b*). However, this difference does not appear to be substantial according to the ROC analysis with a corresponding AUC of 0.620. Obviously, the marked variability of survival times and the overlap of confidence intervals reflects the comparable low AUC in the ROC method.

Conclusion

The examples show that even in case of a significant log rank test, the real difference may be substantial. Particularly, the AUC helps to visualize the difference between groups of time to event data. The ROC method may be useful of additional analysis of survival data. The method seems to be more sensitive than conventional test for separating factors such as the log rank test. An advantage is that the difference between two Kaplan-Meier curves is transferred to a normalized 1 * 1 unit diagram and that the difference is proportional to the AUC of the ROC curve.

References

1. *Altman D.* Practical Statistics for Medical Research. Chapman and Hall. NY 1991
2. *Jensen K, Müller H-H, Schäfer H.* Regional confidence bands for ROC curves. Statistics in Medicine 19: 493-509, 2000
3. *Kalbfleisch JD, Prentice RL.* The Statistical Analysis of Failure Time Data. J. Wiley, NY 1980
4. *Kaplan EL, Meier P.* Nonparametric estimation from incomplete observations. J Am Stat Assoc 53: 457-481, 1958