Pak. J. Statist.
2010 Vol. 26(4), 659-671

EVALUATING CUT-OFF VALUES WITH TIME-DEPENDENT ROC CURVES FOR Ki67

Sigirli, D.[1], **Ercan, I.**[1*], **Jin, Y.**[2], **Paksoy, E.**[3]
Tolunay, S.[4] and **Tasdelen, I.**[3]

[1] Department of Biostatistics, Faculty of Medicine,
Uludag University, Bursa, Turkey.
Email: ercan@uludag.edu.tr (Corresponding Author)

[2] Department of Biostatistics, School of Public Health,
University of Washington, Washington, USA.

[3] Department of Surgery, Yunus Emre State Hospital,
Eskisehir, Turkey.

[4] Department of Pathology, Faculty of Medicine,
Uludag University, Bursa, Turkey.

[5] Department of Surgery, Faculty of Medicine,
Uludag University, Bursa, Turkey.

ABSTRACT

Receiver operating characteristic (ROC) curve is one of the most active research areas in medical statistics. While classic ROC curve analysis is a well-established method to study the accuracies of clinical markers, it may be suboptimal for analyzing outcomes over time, such as prognosis. When the outcome variable of interest is an event that occurs not concurrent with the tests, time-dependent ROC curves should be used to accommodate this time lag.

A commonly used prognostic factor in breast cancer is proliferation rate of tumor cells. Ki67 monoclonal antibody is one of the agents that can be used to evaluate proliferation rate. The goal of this study is to evaluate the cut-off values of Ki67 as a diagnostic marker in breast cancer and to show the difference in terms of diagnostic accuracy using the classical ROC analysis and time-dependent ROC analysis for this marker.

In classical ROC analysis, no significant evidence was found that the area under the curve (AUC) was different from the null hypothesis value 0.5 for both Ki67 in lymph nodes and Ki67 in tumor cells. But in time-dependent ROC analysis that accommodates the event time outcomes, significant tests were found. Using bootstrapped variances, AUCs were significantly larger than 0.5 for 37-136th months and for lymph nodes AUCs were significantly larger than 0.5 for the 23-136th months.

KEYWORDS

ROC curve; Time-dependent ROC; Cut-off; Ki67.

1. INTRODUCTION

To determine the diagnostic accuracy of qualitative and quantitative diagnostic methods, commonly used measures include sensitivity, specificity and the receiver operating characteristic (ROC) curves [1-3].

ROC analysis has widely been used in medical data analysis in measuring discriminatory ability of diagnostic or prognostic tests. This makes the ROC analysis one of the most active research areas in medical statistics. Various methods including parametric and semi-parametric estimation methods have been presented for estimating the ROC curve [4-9].

The area under the curve (AUC) is commonly used as a summary measure of the ROC curve. It indicates the overall performance of a diagnostic test in terms of its accuracy at various diagnostic thresholds used to discriminate cases and non-cases of disease. Since the curve located in the unit square, we have $0 \leq AUC \leq 1$. A test with an area under the ROC curve of 1.0 is perfectly accurate because the sensitivity is 1.0 when the false positive rate is 0.0. In contrast, a test with an area of 0.0 is perfectly inaccurate. The practical lower limit for the AUC of a diagnostic test is 0.5 which reflects completely random classifier [10]. Because AUC is an important index for the ROC curve, the estimation method has been discussed in many studies [11-15].

ROC curve analysis is a well-established method to study the accuracies discriminatory ability of diagnostic or prognostic tests of clinical markers. But in prospective cohort studies the disease status of a subject often changes during the course of the study and there is often a time lag between when the marker is measured and the occurrences of disease [16]. So it may be insufficient for analyzing outcomes over time, such as prognosis. To evaluate the accuracy of such marker, one needs to take the time lag into account since the accuracy may be higher when the markers are measured closer to the onset of disease. When the outcome variable of interest is an event that occurs some time after the test is measured, ROC curves can be considered to depend on time variable. Two approaches have been proposed by Heagerty et al. for being used in this situation [17].

This extension of ROC analysis which handles time-to-event outcomes has been used in identifying high-risk patients with colon cancer and diffuse large B-cell lymphomas, to study gene expression profiles in lung cancer patients and to evaluate diagnostic accuracy of prostate specific antigen in prostate cancer patients [18-20].

Ki67 monoclonal antibody is a nuclear antigen that is present in the mid-G1, S, G2 and the entire M phase of the cell cycle [21]. A commonly used prognostic factor in breast cancer is proliferation rate of tumor cells. Ki67 monoclonal antibody is one of the agents that can be used to evaluate proliferation rate of tumor cells [22, 23].

So far, prognostic impact or predictive value of ki67 as a clinical marker to characterize breast cancer has been examined in many studies [24-27]. Different cut-off determination methods have been used in different studies, most of which are classical ROC analysis [23, 28-30]. In this study, we aimed to evaluate the cut-off values of ki67 with time-dependent ROC analysis and to show the difference between the classical ROC analysis and time-dependent ROC analysis for this marker, and to display the clinical utility and possible benefits of time-dependent ROC curve analysis for determining the prognostic accuracies of the markers.

2. METHODS

Patients treated at the Uludag University Faculty of Medicine Department of Surgery between 1997-2002 were evaluated. Only invasive ductal carcinomas and stage II-III patients included to the study. The study was conducted in accordance with Declaration of Helsinki guidelines. The study protocol, patient information form and informed consent form were approved by the local ethics committees. Informed consent was obtained from all the patients.

Cytological slides from paraffin blocks which are found in Pathology department were prepared. Ki-67 immunostaining was performed on 4 µm thick sections cut from formalin fixed, paraffin wax embedded tissue using the streptavidin–biotin complex peroxidase technique. In each area 1000 nuclei were counted using, 10x40 objective magnification. Ki67 proliferation index was determined as positive nuclei number/1000. Proliferation index were measured both in tumor cells and in lymph node.

Two ROC curve estimators were proposed, both of which can handle time-to-event outcomes, by Patrick Heagerty and colleagues. The first estimator is based on Kaplan-Meier survivor function methods. An alternative estimator that does guarantee monotonicity is based on a nearest neighbor estimator for the bivariate distribution function of (X, T), where T represents survival time which always yields monotone ROC curves and X represents the diagnostic marker [17, 31]. Monotonocity is the condition that sensitivity and specificity are monotone in X and it can be shown with the inequality

$$P\left[X > c \middle| D(t) = 1\right] \geq P\left[X > c' D(t) = 1\right] \tag{2.1}$$

for all c’>c. Here c and c’ are different cut-off points, X is the covariate and D(t) is the event indicator at time t.

If patient i has an event/dies prior to time t, the binary disease/failure variable $D_i(t)$ becomes 1; otherwise, $D_i(t)$ becomes 0. For a diagnostic marker X, both sensitivity and specificity are defined as a function of time t as given follows.

$$\text{Sensitivity (c, t)} = P\left\{X > c \middle| D(t) = 1\right\} \tag{2.2}$$

$$\text{Specificity (c, t)} = P\left\{X \leq c \middle| D(t) = 0\right\} \tag{2.3}$$

The corresponding ROC(t) curve for any time t is defined as the plot of {sensitivity(c, t)} versus {1 - specificity(c, t)}, with cutoff point c varying. X is the covariate and D(t) is the event indicator at time t . The area under the curve, AUC(t), is defined as the area under the ROC(t) curve. A nearest neighbor estimator for the bivariate distribution function is used for estimating these conditional probabilities accounting for possible censoring [17].

First of all, to see the performance of two markers throughout the study period, AUC(t)s obtained for different values of t. We used the nearest neighbor estimator with a span of lambda(n) = 0.05. Second, using bootstrapped variances, we obtained the 95% confidence interval for the AUC by 500-bootstraped replications of the data. We tested whether the area under the curve was different from the value 0.5 (no discriminating power) for both Ki67 in lymph nodes and Ki67 in tumor cells, for different values of t. After that, we determined the significant AUC (t) values and found the optimal cut-off value with the highest Youden J index for each significant AUC (t).

The Youden Index measures the effectiveness of a diagnostic marker and permits the selection of an optimal threshold value or cut-off point for the biomarker of interest. In most cases, there is an inverse relationship between sensitivity and specificity, so moving the cutoff point increases one while reducing the other. Youden J occurs at the cutoff point that optimizes the biomarker's differentiating ability when equal weight is given to sensitivity and specificity. Youden J index was calculated as below over all cut points c; $-\infty < c < \infty$.

$$J = \max imum\left\{sensitivity(c) + specificity(c) - 1\right\} \tag{2.4}$$

The cutoff point that achieves this maximum is referred to as the optimal cutoff point [32-35].

Time-dependent ROC curve analysis was performed with R software, version 2.10 and with the "survivalROC" package [36, 37].

3. RESULTS

Study group consist of 81 female breast cancer patients with the mean ages of 51.14±1.42 (mean ± standard error of mean) years and ranging from 26 to 80. Mean tumor sizes were 3.14 ± 0.24 cm. 55.6% of the patients were in stage 2 and 44.4% were in stage 3. The study ended at 136 months and 81.5% of the patients were still alive at the end of the study. Kaplan-meier survival mean was 114.03 months (standard error: 5.06, 95% CI: 104.13-123.94). Clinical and demographic characteristic of the patient were given in Table 1.

In classical ROC analysis, no significant evidence was found that the area under the curve (AUC) was different from the null hypothesis value 0.5, which indicates no discriminating power, for both Ki67 in lymph nodes and Ki67 in tumor cells.

In time-dependent ROC analysis that accommodates the event time outcomes, significant results were found. For tumor nodes, using bootstrapped variances, AUC was found significantly larger than 0.5 for 37-136th months. No significant AUC was found for the other time points (Table 2).

For lymph nodes AUCs were significantly larger than 0.5 for the 23-136th months. No significant AUC was found for the other time points (Table 3).

ROC curves for significant AUC(t) values are given in figure 1 and 2. It was seen that prognostic performance of lymph ki67 was better than the tumor ki67 marker for t=20, 50, 60 and 70 (Figure 3).

4. DISCUSSION

Although in classical ROC curve analysis, no significant evidence was found that the AUC was different from the null hypothesis value 0.5 which indicates no discriminating power for both ki67 in lymph nodes and ki67 in tumor cells; in time-dependent ROC curve analysis, tumor and lymph ki67 were found as significant diagnostic markers at some time intervals. This is a finding which shows us the importance of using time-dependent approach for censored events.

Different cut-off determination methods have been used in different studies, most of which are classical ROC analysis. While in some of them arbitrary cut-off values have been used, others have used classical ROC curve analysis which does not take into account the censored events [23-30]. In our study, we demonstrated that, cut-off values for ki67 index showed variations by time. As a matter of fact, although we couldn't get significant differentiating performance in classical ROC curve analysis, we got significant results for time-dependent ROC curve analysis after the 37th month for tumor cells, and after the 23rd month for lymph nodes. The impact of ki67 on clinical outcome may therefore only be relevant at later follow-up times.

Cut-off values for ki67 show differences according to different time points. So, in evaluating patients it would be proper to make statements using cut-off values obtained for these times. Our study also keeps its originality in terms of supporting reference to clinical decision makers about ki67 index for tumor cells and lymph nodes.

This approach has several advantages over standard ROC curve approaches. Time-dependent ROC curve analysis may increase the identification of prognostically relevant markers and can visualize the changes in discriminatory power of the marker from the time of diagnosis over various follow-up times. In addition, it provides information on the time interval over which the marker is most reliable, how clinically relevant it is and which cut-off scores should be used to discriminate between better or worse survival [18].

Several approaches have been proposed for deriving ROC curves for time-to-event data. To take into account the covariate effects, methodologies which use regression analysis, semiparametric estimation methods or Cox model have been proposed [16, 38-39]. Pepe et al. have considered marker performance for outcomes that are not simply binary but that are event time outcomes and given existing methods and discussed issues that lead to preference for one method over another [40].

So when the status of patients contains censoring, estimates ignoring censoring can be biased. Therefore, time-dependent ROC analysis can be performed to correct for the bias and integrate the time dimension.

REFERENCES

1. McNeil, B.J. And Hanley J. (1984). Statistical Approaches to the Analysis of ROC curves, *Med Decis Making*, 4(2), 136-149.
2. Pepe M.S. (2003). *The statistical evaluation of medical tests for classification and prediction*. Oxford. New York.
3. Zhou, X.H., Obuchowski, N.A., McClish, D.K. (2002). *Statistical methods in diagnostic medicine*. Wiley & Sons Interscience. New York.
4. Swets, J.A. and Picett, R.M. (1982). *Evaluation of Diagnostic Systems: Methods from Signal Detection theory*. Academic. New York.
5. Hanley, J.A. (1989). Receiver Operating Characteristic (ROC) Methodology: the state of the art. *Clinical Reviews in Diagnostic Imaging*, 29, 307-335.
6. Begg, C.B. (1991). Advances in statistical methodology for diagnostic medicine in the 1980s. *Statistics in Medicine*, 10, 1887-1895.
7. Wieang, S., Gail, M.H., James, B.R. and James, K.L. (1989). A family of nonparametric statistics for comparing diagnostic markers with paired or unpaired data. *Biometrika*, 76, 585-592.
8. Wan, S. and Zhang, B. (2007). Smooth semiparametric receiver operating characteristic curves for continuous diagnostic tests. *Statistics in Medicine*, 26, 2565-2586.
9. Zhang, B. (2006). Semiparametric hypothesis testing procedure for the ROC curve area under a density ratio model. *Computational Statistics & Data Analysis*, 50(7), 1855-1876.
10. Obuchowski, N.A. (2003). Receiver operating characteristic curves and their use in radiology. *Radiology*, 229, 3-8.
11. Hanley, J.A. and McNeil, B.J. (1982). The meaning and use of the area under an ROC curve. *Radiology,* 143(1), 29-36.
12. Bamber, D.C. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *J. Math.Psy.*, 12, 387-415.
13. DeLong, E.R., DeLong, D.M. and Clarke-Pearson, D.L. (1988). Comparing the areas under two ormore correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics,* 44, 837-845.
14. Brownie, C., Simonoff, J.S., Hochberg, Y. and Reiser, B. (1988). Estimating Pr(X<Y) in categorized data using ROC analysis. *Biometrics*, 44, 615-621.
15. Qin, G. and Zhou, X.H. (2006). Empirical likelihood inference for the area under the ROC curve. *Biometrics,* 62, 613-622.
16. Cai, T., Pepe, M.S., Zheng, Y., Lumley, T. and Jenny, N.S. (2006). The sensitivity and specificity of markers for event times. *Biostatistics,* 7, 182-197.
17. Heagerty, P.J., Lumley, T. and Pepe, M.S. (2000). Time-Dependent ROC Curves for Censored Survival Data and a Diagnostic Marker. *Biometrics,* 56(2), 337-344.
18. Adams, H., Tzankov, A., Lugli A. and Zlobec I. (2009). New time-dependent approach to analyse the prognostic significance of immunohistochemical biomarkers in colon cancer and diffuse large B-cell lymphoma. *J Clin Pathol*, 62,986-997.
19. Etzioni, R., Pepe M.S., Longton, G., Hu C. and Goodman, G. (1999). Incorporating the Time Dimension in Receiver Operating Characteristic Curves: A Case Study of Prostate Cancer. *Med. Decis. Making*, 19, 242-251.

20. Guo, L., Ma, Y., Ward, R., Castranova, V., Shi, X. and Qian, Y. (2006). Constructing Molecular Classifiers for the Accurate Prognosis of Lung Adenocarcinoma. *Clinical Cancer Research*, 12, 3344.
21. Gerdes, J., Lemke, H., Baisch, H., Wacker, H.H., Schwab, U. and Stein, H.J. (1984). Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. *Immunol,* 133(4), 1710-1715.
22. Gasparini, G., Boracchi, P., Verderio, P. and Bevilacqua, P. (1994). Cell kinetics in human breast cancer: comparison between the prognostic value of the cytofluorimetric S-phase fraction and that of the antibodies to Ki-67 and PCNA antigens detected by immunocytochemistry. *Int. J. Cancer,* 57, 822-829.
23. Ahlin, C., Aaltonen, K., Amini, R.M, Nevanlinna, H., Fjallskog, M.L. and Blomqvist, C. (2007). Ki67 and cyclin A as prognostic factors in early breast cancer. What are the optimal cut-off values? *Histopathology*, 51, 491-498.
24. Cheang, M.C.U., Chia, S.K., Voduc, D., Gao, D., Leung, S., Snider, J., Watson, M., Davies, S., Bernard, P.S., Parker, J.S., Perou, C.M., Ellis, M.J. and Nielsen, T.O. (2009). Ki67 Index, HER2 Status, and Prognosis of Patients with Luminal B Breast Cancer. *J. National Cancer Institute*, 101(10), 736-750.
25. Penault-Llorcaac, F., Abriala, C., Raoelfilsa, I., Cholletac, P., Cayrea, A., Mouret-Reyniera, M.A., Thivata, E., Mishellanya, F., Gimberguesa, P. and Durando, X. (2008). Changes and Predictive and Prognostic Value of the Mitotic Index, Ki-67, Cyclin D1, and Cyclo-oxygenase-2 in 710 Operable Breast Cancer Patients Treated with Neoadjuvant Chemotherapy. *The Oncologist,* 13(12), 1235-1245.
26. Buxant, F., Anaf, V., Simon, P., Fayt, I. and Noël, J.C. (2002). Ki-67 Immunostaining Activity is Higher in Positive Axillary Lymph Nodes than in the Primary Breast Tumor, *Breast Cancer Research and Treatment,* 75(1), 1-3.
27. Pinto, A.E., André, S., Pereira, T., Nóbrega, S. and Soares, J. (2001). Prognostic comparative study of S-phase fraction and Ki-67 index in breast carcinoma. *J. Clin. Pathol.*, 54, 543-549.
28. Haapasalo, J., Mennander, A., Helen, P., Haapasalo, H. and Isola, J. (2005). Ultrarapid Ki-67 immunostaining in frozen section interpretation of gliomas, *J. Clin. Pathol.,* 58(3), 263-268.
29. Park, D., Kåresen, R., Noren T. and Sauer, T. (2007). Ki-67 expression in primary breast carcinomas and their axillary lymph node metastases: clinical implications. *Virchows ArchivVolume,* 451(1), 11-18.
30. Mohsenifar, J., Almassi-Aghdam, M., Mohammad-Taheri, Z., Zare, K., Jafari, B., Atri, M., Mortazavi, S.H., Azadeh, P., Bagherzadeh, M., Azargashb, E. and Rahimi, F. (2007). Prognostic Values of Proliferative Markers ki-67 and repp86 in Breast Cancer. *Arch Iranian Med*, 10(1), 27-31.
31. Akritas, M.J. (1994). Nearest Neighbor Estimation of a Bivariate Distribution under Random Censoring. *Ann. Statist.*, 22, 1299-1327.
32. Youden, W.J. (1950). Index for rating diagnostic tests. *Cancer*, 3, 32-35.
33. Faraggi, D. (2000). The effect of random measurement error on receiver operating characteristic (ROC) curves. *Statistics in Medicine,* 19(1), 61-70.
34. Perkins, N.J. and Schisterman, E.F. (2005). The Youden Index and the Optimal Cut-Point Corrected for Measurement Error. *Biomtrc J.,* 47(4), 428-441.

35. Schisterman, E.F., Faraggi, D., Reiser, B. and Hu, J. (2008). Youden Index and the optimal threshold for markers with mass at zero. *Statist. Medicine*, 27(2), 297-315.
36. R Development Core Team (2009). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna. Austria.
37. Heagerty, P.J., packaging by Saha, P. (2008). R Package: SurvivalROC. CRAN.
38. Heagerty, P.J. and Zheng, Y. (2005). Survival model predictive accuracy and ROC curves. *Biometrics*, 61, 92-105.
39. Song, X. and Zhou, X.H. (2008). A semiparametric approach for the covariate specific ROC curve with survival outcome. *Statistica Sinca*, 18, 947-965.
40. Pepe, M., Zheng, Y. and Jin, Y. (2008). Evaluating the ROC Performance of Markers for Future Events. *Lifetime Data Anal*., 14, 86-113.

TABLES

Table 1
Summary of characteristics of the study population (n = 81)

Characteristics		**n (%) or mean±sem**
Age (years)		51.14 ±1.42
Primary tumor size (cm)		3.14± 0.24
Size of metastatic lymph node (cm)		14.98±0.80
Kaplan-meier survival mean (month)		114.03±5.06
Histologic grade	Grade 1	8 (9.9%)
	Grade 2	45 (55.6%)
	Grade 3	28 (34.6%)
Progesterone receptor	Negative	38 (46.9%)
	Positive	43 (53.1%)
Oestrogen receptor	Negative	33 (%34,7)
	Positive	62 (%65.3)
Nuclear Grade	Grade 1	9 (11.1%)
	Grade 2	42 (51.9%)
	Grade 3	30 (37%)
TNM stage group	Stage II	45 (55.6%)
	Stage III	36 (44.4%)

sem: Standard error of mean

Table 2
Time-dependent ROC curve analysis results and accuracy summaries for tumor ki67.

Time Interval (month)	AUC	p	cut-off	Youden J	Sensitivity	Specificity	Positive Likelihood Ratio	Negative Likelihood Ratio
1-4	0.006	1	-	-	-	-	-	-
5-17	0.425	0.657	-	-	-	-	-	-
18	0.723	0.227	-	-	-	-	-	-
19	0.547	0.419	-	-	-	-	-	-
20-22	0.637	0.239	-	-	-	-	-	-
23-26	0.640	0.175	-	-	-	-	-	-
27-34	0.652	0.111	-	-	-	-	-	-
35	0.646	0.083	-	-	-	-	-	-
36	0.644	0.063	-	-	-	-	-	-
37-45	0.651	0.039	125	0.317	0.901	0.415	1.540	0.239
46-59	0.635	0.046	125	0.331	0.909	0.422	1.573	0.217
60	0.652	0.022	125	0.349	0.918	0.431	1.613	0.190
61-68	0.636	0.031	125	0.354	0.917	0.438	1.632	0.189
69-136	0.641	0.020	141	0.383	0.911	0.472	1.725	0.189

Table 3
Time-dependent ROC curve analysis results and accuracy summaries for lymph ki67

Time Interval (month)	AUC	p	cut-off	Youden J	Sensitivity	Specificity	Positive Likelihood Ratio	Negative Likelihood Ratio
1-4	0.006	1	-	-	-	-	-	-
5-17	0.313	0.900	-	-	-	-	-	-
18	0.705	0.255	-	-	-	-	-	-
19	0.674	0.211	-	-	-	-	-	-
20-22	0.730	0.086	-	-	-	-	-	-
23-26	0.735	0.026	259	0.392	0.832	0.560	1.891	0.300
27-34	0.721	0.023	259	0.423	0.856	0.480	1.646	0.300
35	0.742	0.003	259	0.456	0.879	0.577	2.078	0.210
36	0.725	0.004	257	0.449	0.879	0.570	2.044	0.212
37-45	0.748	0.001	259	0.495	0.893	0.602	2.244	0.178
46-59	0.703	0.006	259	0.415	0.819	0.596	2.027	0.304
60	0.702	0.005	259	0.441	0.835	0.606	2.119	0.272
61-68	0.685	0.008	250	0.434	0.846	0.589	2.058	0.261
69-136	0.631	0.042	227	0.283	0.759	0.524	1.595	0.460

FIGURES

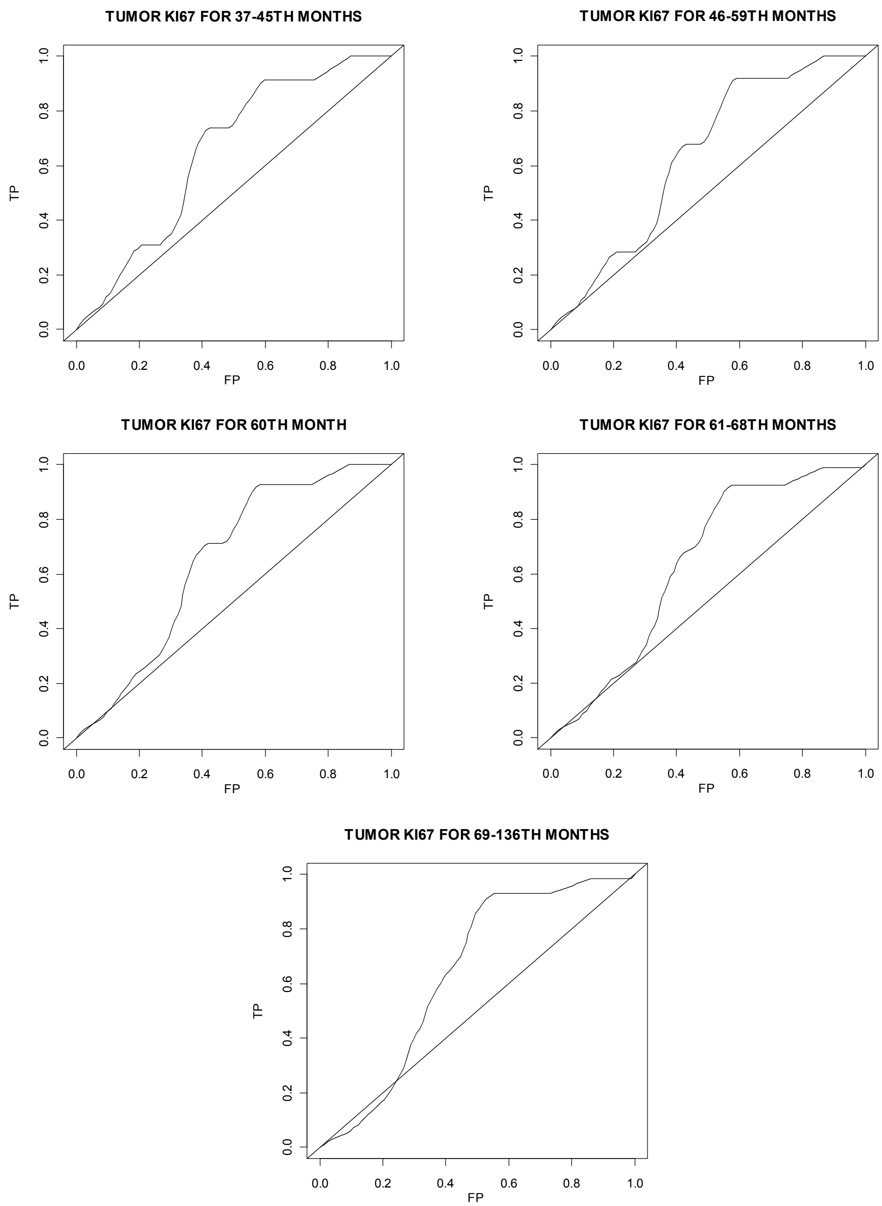

Fig. 1: ROC curves for tumor ki67 marker.

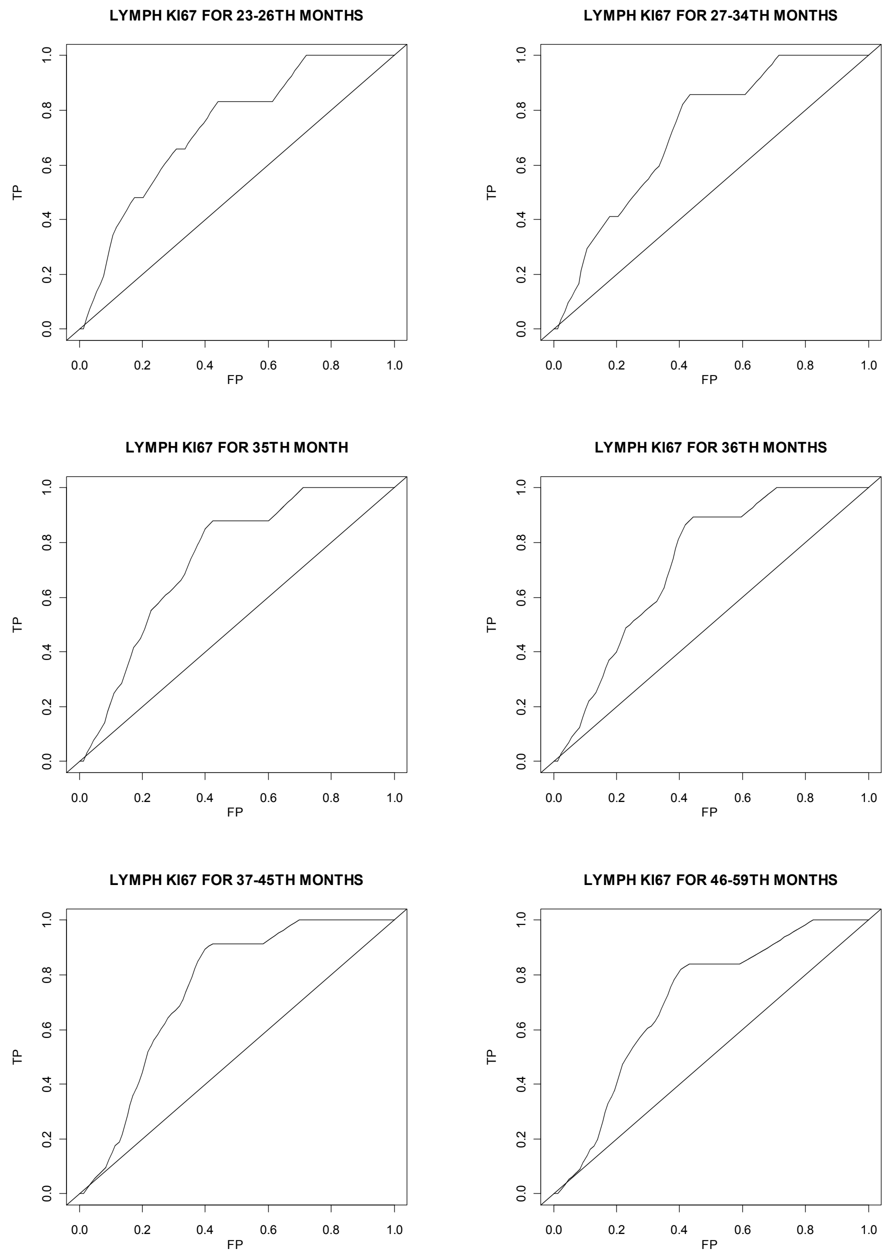

Fig. 2: ROC curves for lymph ki67 marker

Contd…. Next page

Fig. 2: (continued)

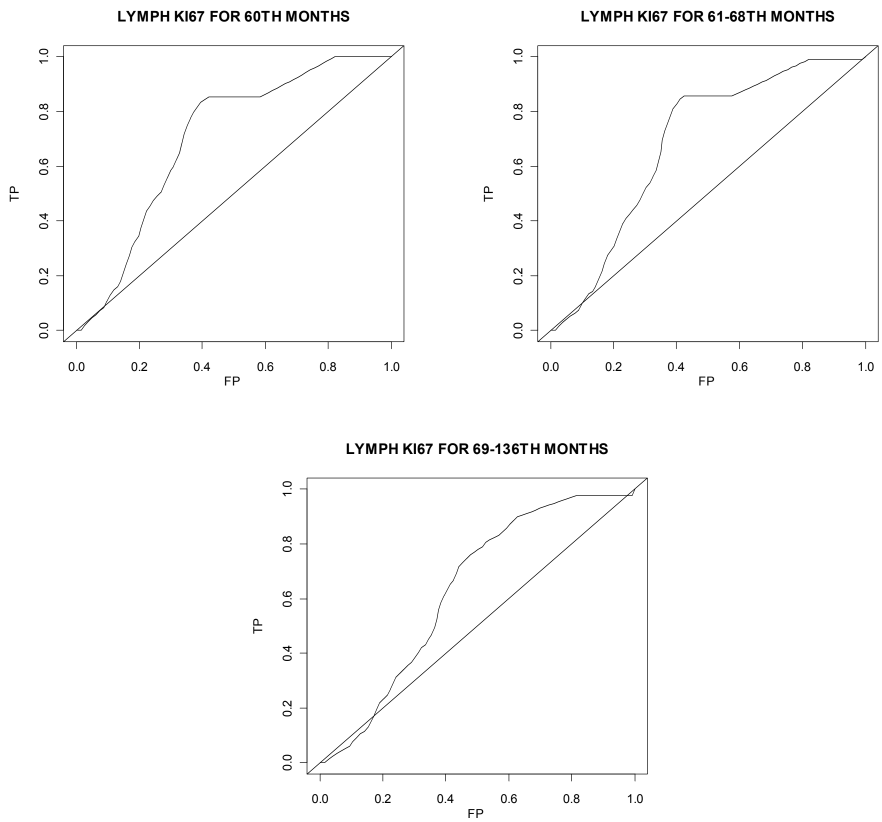

Fig. 2: ROC curves for lymph ki67 marker.

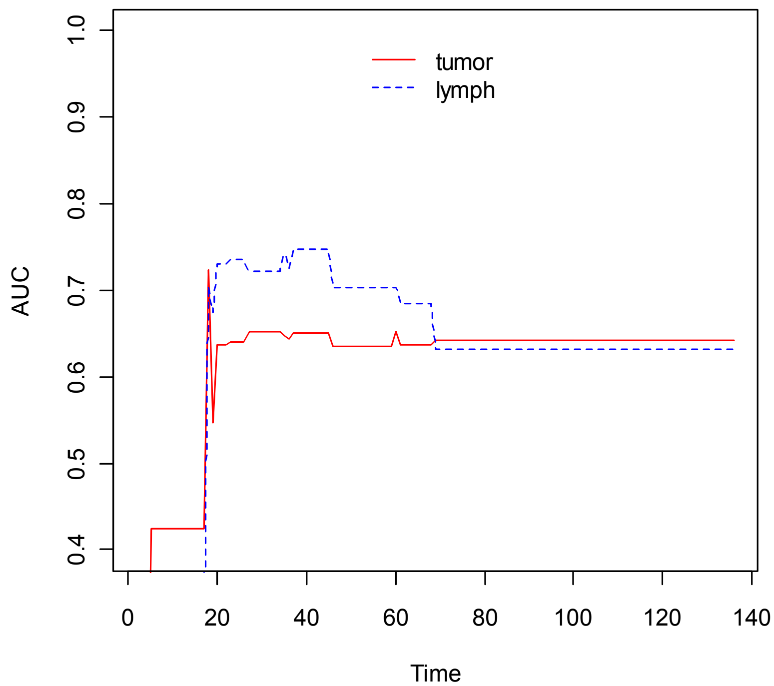

Fig. 3: AUCs for tumor and lymph ki67 at different time points.