Araştırma Makalesi / *Research Article*

doi: 10.5505/sakaryamj.2015.95914

Covariate Adjusted ROC Curve Analysis and An Application

Ortak Değişkene Göre Düzeltilmiş ROC Eğrisi Yöntemi ve Bir Uygulama

Ünal Erkorkmaz[1], Ertuğrul Çolak[2], Cengiz Bal[2], Kazım Özdamar[2], İlker Etikan[3], Hasan Çetin Ekerbiçer[4]

[1] Sakarya Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Sakarya, Türkiye

[2] Eskişehir Osmangazi Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Eskişehir, Türkiye

[3] Yakın Doğu Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Lefkoşa, Kuzey Kıbrıs Türk Cumhuriyeti

[4] Sakarya Üniversitesi Tıp Fakültesi Halk Sağlığı Anabilim Dalı, Sakarya, Türkiye

Abstract

Received: 03.02.2015 **Accepted:** 20.02.2015

Objective: Aim of this study is to analyze the change of the area under the adjusted ROC (AdjROC) curve in certain conditions via binormal distribution model using simulation studies and application of this algorithm to real data. Materials and Methods: Data sets simulated according to various conditions. PSA and age values of 125 patients who were examined prostate biopsy with pre-diagnosis of prostate cancer in Gaziosmanpasa University Faculty of Medicine Department of Urology at the years of 2005 to 2007. An algorithm and code program was written that make simulation according to various condition using PROC IML procedure in SAS statistical software.Results: According to the simulation study, if biomarker indicators in healthy group are constant and are lower or equal in healthy group than/to disease group, both adjusted AUC (AdjAUC) and AUC have small values and, no significant difference was found between them. The AUC was significantly larger when the biomarker indicators in disease group were higher. In addition, if the correlation between the covariate and biomarker is high in disease group and if AUC is approximately 0.75, then there is significant difference between adjusted AUC and AUC. PSA (Prostate Specific Antigen), a biomarker used for prostate cancer diagnosis, was analyzed based on the adjustments by age. It was found that adjusted AUC value was higher than unadjusted AUC value. Conclusions: For the adjusted ROC model being applicable, covariate and biomarker distributions must show double binormal distribution. If the biomarker can distinguish disease and healthy individuals correctly, then covariate is not needed. If correlation of healthy is approaching to 0 and correlation of disease is 0.50, and if AUC is less than 0.75, then covariate must be included in the model. Model does not work well when sample size of disease and healthy are less than 50.

Keywords: Adjusted ROC, AUC, Covariate, PSA, Simulation

Özet

Geliş Tarihi : *03.02.2015* ***Kabul Tarihi:*** *20.02.2015*

Amaç: Bu araştırmada, benzetim çalışmalarından yararlanarak düzeltilmiş ROC eğrisi altında kalan alanın belirli koşullardaki değişiminin iki değişkenli normal dağılım modeli ile incelenmesi ve bu algoritmanın gerçek verilerle uygulanması amaçlanmıştır. Gereç ve Yöntemler: Benzetimde kullanılacak veri seti farklı koşullar altında türetilmiştir. Gerçek uygulama verisi olarak Gaziosmanpaşa Üniversitesi Tıp Fakültesi Üroloji Anabilim Dalında 2005-2007 yılları arasında prostat kanseri ön tanısı için prostat biyopsisi yapılan 125 hastanın PSA değerleri ile yaşları kullanılmıştır. Algoritma ve kodlar farklı koşullardaki benzetim modellerine göre SAS istatistik yazılımında PROC IML prosedürü kullanılarak yazılmıştır. Bulgular: Benzetim çalışmasına göre, biomarker göstergeleri sağlam grupta sabit ve hasta grupta sağlam gruba göre daha düşük veya eşit ise hem AUC (ROC Eğrisi Altında Kalan Alan) hem de düzeltilmiş AUC'nin düşük değerler aldığı bulunmuş ancak aralarında önemli fark görülmemiştir. Hasta grupta daha yüksek biomarker göstergeleri olduğunda ROC eğrisi altında kalan alan belirgin şekilde yüksek bulunmuştur. Ayrıca biomarker ile ortak değişken arasındaki korelasyon hasta grupta yüksek ve AUC yaklaşık 0.75 ise düzeltilmiş AUC ile AUC arasındaki fark önemli bulunmuştur. Prostat Kanseri biomarker'ı olan PSA'yı yaşa göre düzeltilmiş olarak incelediğimizde, düzeltilmiş AUC değerinin düzeltilmemiş AUC değerine göre daha yüksek olduğu bulunmuştur.

Sonuç: Düzeltilmiş ROC modelinin uygulanabilir olması için ortak değişken ile biomarker, dağılımlarının çift iki değişkenli normal dağılım göstermesi gerekmektedir. Biomarker, hasta ve sağlam ayrımını iyi yapıyorsa ek bir değişkene ihtiyaç duyulmamaktadır. Sağlam gruptaki korelasyon 0'a yaklaştıkça ve hasta gruptaki korelasyon 0.50 ise ve AUC 0.75 ve daha küçük ise ortak değişkenin modele katılması gerekir. Hasta ve sağlam gruplarda örnek büyüklüğü 50'den küçük olması durumunda model etkili biçimde çalışmamaktadır.

Anahtar Kelimeler: *Düzeltilmiş ROC, AUC, Ortak Değişken, PSA, Benzetim.*

Yazışma Adresi / Correspondence: Ünal Erkorkmaz,
Sakarya Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, Korucuk Kampüsü 54290 Adapazarı-Sakarya, Türkiye
E-mail: uerkorkmaz@gmail.com

Introduction

In medicine, laboratory tests are substantially utilized when diagnosing diseases. Today, there are numerous tests that determine hematological, biochemical, and histopathological properties of individuals. When diagnosing diseases, the values obtained from these tests are the most important source of reference of the doctor in addition to the radiological imaging, physical and interventional examination findings. The results of the laboratory tests (Y, biomarker) that reveal the biological properties of individuals cannot be directly interpreted as the evidence of disease. It must be exactly clarified that in larger or smaller than which values, biomarkers point the disease. For biomarkers to be used in healthy-disease discrimination, appropriate cutting points must be determined in a valid and reliable way.[1-4]

ROC (Receiver Operating Characteristics Curve) analysis is a very commonly used method in determining which values of the numerical results ($Y \geq k$, $Y < k$) that are obtained from laboratory tests that helps diagnosing a disease, interventional results or physical examination, points the existence of a disease or a phenomenon in a valid and a reliable way.[4]

In the case where Y is a measurable variable for all the healthy and disease individuals in the society, the threshold value that will be used for distinguishing disease from healthy individuals must be a value that will minimize false positive and false negative results in diagnosis of X disease that is examined. For Y to have a distribution with similar parameters in healthy and disease group, makes it harder to use Y values in diagnosis. Furthermore, the existence of covariates ($Z_1, Z_2, \ldots, Z_p$) that has a change together with Y significantly affects the validity and consistency of the decisions.

Risk factors of most of the diseases include factors such as age, sex, occupation, race, BMI, total cholesterol, LDL, HDL, WBC, residence, time of affection, dose, daily activities and covariates (Z_j). These factors and the values of covariates have effect on the biomarkers that are obtained from laboratory tests.[3] When effects of covariates on the biomarker increases, wrong evaluation of disease as being healthy or wrong evaluation of healthy individuals as being disease is of concern. The use of covariates according to the corrected value instead of using laboratory results directly is of great importance in recent years. The ROC analysis that is known since 1950s, initiated the studies about adding covariate values into analysis in using biomarkers in diagnosis and treatment. In most of these researches, it is concluded that the corrections in minimizing the false positive and false negative results in accordance with the covariates, will be active in determining the threshold values and will improve the performance of diagnosis test results.[1-3, 5-17]

Alonzo et al. developed the ordinal regression model with hidden variables to correct the ordinal test results and to add the effects of covariates into the evaluation.[5] However, this approach is not considered applicable for continuous biomarkers.

Pepe, proposed the semi-parametric ROC approach to obtain ROC curves that are corrected for covariates.[12] In ROC analysis, Pepe, claim that analysis can be made by considering the effects of covariates through generalized linear model.[13] Alonzo et al. proposed a parametric regression model for ROC curve and showed that model is valid and reliable with simulation studies.[5]

Cai and Cai *et al*. have generalized parametric ROC regression model.[18,6,19]

Punglia *et al*., made adjustments by age when reviewing distinctive performance of the PSA (prostate specific antigen) that is the prostate cancer diagnostic test, and found that when compared with unrevised analysis, the adjustment that is corrected according to covariates, significantly increases the area under the ROC curve in PSA test.[20] Ghosh et al. applied the covariate-adjusted regression model on the data obtained from molecular identification study of prostate cancer in their studies.[8]

Zhang *et al*. stated that linear regression approach can be used for comparison of ROC curves and the effects of covariate can also be included into the model in their work.[21]

In the studies of Schisterman *et al*., it is claimed that the area under the ROC curve (AUC) can be combined with a good

linear combination where it is maximized within all possible linear combinations when evaluating performance of diagnostic tests that are affected by covariates and ROC curve can be estimated by the help of this linear combination.[15] Also the performances of the diagnosis test of two coronary heart diseases are compared by taking into account the effects of two covariates, such as age and gender. In another study, Schisterman *et al*. developed a flexible model alternative to the standard model they proposed previously and compared it with this model. [22, 15]

The Adjusted ROC (AdjROC) model that is the adjusted measure of the classification accuracy of the diagnosis tests according to the covariates is proposed by Janes *et al*.[2]. AdjROC is the ROC curve that uses thresholds specific to the covariate to define "test positive". In the studies of Janes *et al*. AdjROC is compared with traditional ROC curves, non-parametric or semi-parametric estimators are proposed for AdjROC and asymptotic distribution theory is developed for these estimators.[2] In this study, simulation approaches show that AdjROC estimators perform quite well for small examples.

Between the factors that affect threshold values in ROC analysis; number of units (n_D, n_C) of group of disease and healthy individuals that are analyzed, the distribution of Y in the group of disease and healthy individuals and its parametric values, the position of these distributions, for the scale parameters of the distribution functions to be close to each other and their intricate are the important factors. The studies that take all these factors into consideration are not seen often in resources. In trial simulation studies it is observed that positive or negative correlation of Z's with the Y's that are in the disease and healthy group and the size of these correlations cause significant change on the AUC. It is observed that the number of units in the group of disease and healthy individuals that will be analyzed, the distribution of measured YS, the difference of parametric values in the groups, and the correlation values of Z and Y within each group affects AUC significantly. As it is known, the most important factor that affects AUC is the threshold value. Combined distribution of Y according to the groups and the difference of the groups within Y distributions affect threshold value and therefore the size of AUC.[4,23] Therefore, the activity of the method must be investigated according to the change of the unit numbers of the group of disease and healthy individuals, the difference of the parametric values of Y, the difference of the parametric values of covariates (Z_1, Z_2, … , Z_p) and the size and the direction of the correlation between Y and Z within each group.
In this study;

1. According to the results of simulation studies, by making use of AUC which is important evaluation criteria in ROC analysis, the following items are aimed by the help of the results of AdjROC which is used to determine threshold values of covariate adjusted Y's:
   - a. Investigating the alteration of AUC according to the sample size of similarly, different, less, and large number of disease and healthy groups (n_D, n_C).
   - b. Investigating the alteration of AUC according to dissimilarity of biomarker values (Y) of disease and healthy groups and parameters (μYD, σYD; μYC, σYC; μZD, σZD; μZC, σZC) of covariate (Z).
   - c. Investigating the alteration of AUC according to correlation levels (ρD, ρC) between covariate values and bio-marker values separately for disease and healthy groups.
2. Developing a program devoted to make the AdjROC analysis in the SAS software and getting it used for simulation and data analysis purposes.
3. By the help of ROC and AdjROC, investigating the PSA values that are adjusted by the ages of 125 prostate patients who contacted Gaziosmanpaşa University Medical Faculty Urology Clinic, whose PSA values measured and prostate biopsies done, comparing the AUC levels according to the both of the methods, and exposing the right classification criteria of the clinic distinctions.

Materİals and Methods

In this study, two types of data structure are used:

1. Data sets simulated according to various conditions,
2. PSA and age values of 125 patients who were examined prostate biopsy with pre-diagnosis of prostate cancer in Gaziosmanpasa University Faculty of Medicine Department of Urology at the years of 2005 to 2007.

Simulation Study

An algorithm and code program was written that make simulation according to various condition using PROC IML procedure in SAS statistical software.

In this program, data were simulated according to all combinations of n_D and n_C in the case where different disease sample size changes within the range (n_D=1000, 500, 250, 100, 50, 25 and 10) and healthy sample size changes within the range (n_C =2000, 1000, 500, 250, 100, 50, 25 and 10) in each trial with k=1000 repeated trials.

In the simulation, firstly diagnostic test and covariate data were simulated being separate for disease and healthy individuals. In the simulation, two types of sample unit numbers were chosen being balanced and unbalanced. In the balanced simulation, sample sizes were chosen as 1000, 500, 250, 100, 50, 25 and 10 ensuring n_D=n_C. In the unbalanced simulation, [n_D, n_C] were chosen as [1000,2000], [500,1000], [250,500], [100,200], [50,100], [25,50] and [10,20] ensuring that the ratio of (n_D:n_C) is 1:2.

The correlation between covariate (Z) and bio-marker (Y) data was chosen in 22 different combinations such as [0.00,0.00], [0.00, 0.95], [0.95,0.00], [0.10, 0.10], [0.25,0.25], [0.50,0.50], [0.75,0.75], [0.95,0.95], [0.10,0.25], [0.10,0.50], [0.10,0.75], [0.10,0.95], [0.25,0.10], [0.50,0.10], [0.75,0.10], [0.95,0.10], [0.95,0.25], [0.95,0.50], [0.95,0.75], [0.25,0.95], [0.50,0.95] and [0.75,0.95] for the disease and healthy groups [ρD, ρC]. Data that has a normal distribution with parameter YC, ZC∼N(0,1) were derived for Y and Z variables that are in the healthy group. For patient group, data that has a normal distribution with parameter ZD∼N(0,1) were chosen for the Z variable ZD∼N(μZD, σZD). For Y variable, data sets were derived from normal distribution that has parameters YD∼N(0.5,1), YD∼N(1.0,1), YD∼N(1,1), YD∼N(1.6,1), YD∼N(1.7,1), YD∼N(1.8,1), YD∼N(1.9,1) and YD∼N(2,1), having different average and unit variance than the normal distribution that has parameter YD∼N(μYD, σYD).

Double binormal distribution assumptions were used for conversions in the disease and healthy groups being suitable for analysis model that is used for data analysis.

$$\begin{pmatrix} Y_D \\ Z_D \end{pmatrix} \sim BVN\left(\begin{pmatrix} \mu_{Y_D} \\ \mu_{Z_D} \end{pmatrix}, \begin{pmatrix} \sigma^2_{Y_D} & \sigma_{Y_D}\sigma_{Z_D}\rho_D \\ \sigma_{Y_D}\sigma_{Z_D}\rho_D & \sigma^2_{Z_D} \end{pmatrix}\right)$$

$$\begin{pmatrix} Y_C \\ Z_C \end{pmatrix} \sim BVN\left(\begin{pmatrix} \mu_{Y_C} \\ \mu_{Z_C} \end{pmatrix}, \begin{pmatrix} \sigma^2_{Y_C} & \sigma_{Y_C}\sigma_{Z_C}\rho_C \\ \sigma_{Y_C}\sigma_{Z_C}\rho_C & \sigma^2_{Z_C} \end{pmatrix}\right)$$

In these conversions, correlation coefficients (ρD and ρC) between Y and Z in disease and healthy groups were used. AUC and AdjAUC calculations were based on the model of Janes et al. [9].

The following equations were used for conversion operations and AUC and AdjAUC calculations:

$$\text{Assuming } \frac{\sigma_{Y_D}}{\sigma_{Z_D}} = \frac{\sigma_{Y_C}}{\sigma_{Z_C}} \equiv w \text{ and } AUC = \Phi\left(\frac{\mu_{Z_D}}{\sqrt{\sigma^2_{Z_D} + \sigma^2_{Z_C}}}\right)$$

Unadjusted AUC is,

$$AUC = \Phi\left(\frac{\mu_{Z_D}/\sigma_{Z_D} - \rho_D\mu_{Z_D}w/\sigma_{Y_D}}{\sqrt{1+\rho^2_C + \sigma^2_{Y_C}/\sigma^2_{Y_D} - (1-\rho^2_C)}}\right) \quad (2)$$

in this format, and,
Adjusted AUC (AdjAUC) is,

$$AdjAUC = \Phi\left(\frac{\mu_{Y_D}/\sigma_{Y_D} - \rho_C\mu_{Z_D}/\sigma_{Z_D}}{\sqrt{1+\rho^2_C + (\rho_C-\rho_D)^2\,\sigma^2_{Y_C}/\sigma^2_{Y_D} - (1-\rho^2_C)}}\right) \quad (3)$$

calculated in this format.[1-3]

General parameter assumptions are done from k-times repeated values by including above simulations and conversions, and the codes covering AUC and AdjAUC calculations into a loop for each SAS codes being k=1000. The differences between AUC and AdjAUC values are tested with normal distribution approach by the help of obtained asymptotic parame-

Adjusted ROC Application with PCA Data Set

Because of the binormal distribution assumption in adjusted ROC analysis, it is found that PSA values don't have normal distribution in normal distribution applicability test. So, it is normalized by making logarithmic conversion of PSA values and ROC analysis was applied with data that has been converted.

Results

Results of Simulation Study

As seen in table 1, it is observed that AdjAUC values are significantly greater than the AUC values for data simulated in $n_D=n_C=1000$ conditions and in μYD≥1.50 and ρC=ρD≥0.50 conditions. Besides this, when keeping ρC=0.95 constant in the condition where ρD≥0.50 and keeping ρD=0.95 constant in the condition ρC≥0.25, it is found that AdjAUC values are significantly greater than AUC values. In [ρC=ρD=0]; [ρC=0, ρD=0.95] and [ρC=0.95,ρD=0] combinations, AdjAUC and AUC values have similar values. It is observed that, in general, the value of μYD (The parameter of Y values of disease group) has an effect on AUC. When μYD=0.50 it is observed that both AdjAUC and AUC have low values and AdjAUC and AUC values cannot make a good distinction for disease-healthy events. When μYD=1.00, it observed that it has a distinction greater than of μYD=0.50, but still has not effective distinction. When μYD has a value greater than or equal to 1.50, the area under the curve gets larger and for μYD=2.00, it reaches the highest level. If covariate parameters increase as biomarker parameters, the distinction of disease and healthy individuals increases significantly.

Table 1. The results of simulated data according to the conditions of $n_D=n_C=1000$

μYD	ρC	ρD	AUC	SEMAUC	AdjAUC	SEMAdjAUC	p
1.50	0.25	0.25	0.812	0.008	0.820	0.008	0.316
	0.50	0.50	0.760	0.010	0.793	0.010	0.001
	0.75	0.75	0.702	0.013	0.789	0.016	<0.001
	0.95	0.95	0.652	0.015	0.893	0.023	<0.001
	0.95	0.25	0.652	0.015	0.673	0.017	0.201
	0.25	0.95	0.812	0.008	0.844	0.008	<0.001
1.70	0.25	0.25	0.847	0.007	0.855	0.007	0.287
	0.50	0.50	0.802	0.009	0.836	0.009	<0.001
	0.75	0.75	0.749	0.014	0.844	0.014	<0.001
	0.95	0.95	0.702	0.015	0.954	0.013	<0.001
	0.95	0.25	0.703	0.014	0.729	0.016	0.084
	0.25	0.95	0.847	0.007	0.880	0.007	<0.001
1.90	0.25	0.25	0.878	0.006	0.886	0.006	0.224
	0.50	0.50	0.839	0.009	0.873	0.008	<0.001
	0.75	0.75	0.792	0.011	0.890	0.011	<0.001
	0.95	0.95	0.749	0.014	0.983	0.006	<0.001
	0.95	0.25	0.750	0.015	0.780	0.016	0.052
	0.25	0.95	0.878	0.006	0.909	0.006	<0.001
2.00	0.25	0.25	0.892	0.006	0.899	0.005	0.223
	0.50	0.50	0.855	0.008	0.889	0.007	<0.001
	0.75	0.75	0.811	0.010	0.908	0.009	<0.001
	0.95	0.95	0.770	0.013	0.990	0.003	<0.001
	0.95	0.25	0.771	0.013	0.782	0.014	0.419
	0.25	0.95	0.891	0.006	0.921	0.005	<0.001

As seen in table 2, it is observed that AdjAUC values are significantly greater than the AUC values for data simulated in nD=nC=250 conditions and in μYD=1.50, μYD=1.60, μYD=1.70 and ρC=ρD≥0.75 conditions. Besides this, when keeping ρC=0.95 constant in the condition where ρD≥0.75, and keeping ρC=0.95 constant in the condition where ρC≥0.25, it is found that AdjAUC values are significantly greater than AUC values. In μYD=1.80, μYD=1.90 and ρC=ρD≥0.50 conditions, it is found that AdjAUC values are significantly greater than AUC values. Besides this, when keeping ρC=0.95 constant in the condition where ρD≥0.50 and keeping ρD=0.95 constant in the condition where ρC≥0.25, it is found that AdjAUC values are significantly greater than AUC values. In μYD=2.00 and ρC=ρD=0.25 conditions, it is observed that AdjAUC values are significantly greater than AUC values. When keeping ρC=0.95 constant in the condition where ρD≥0.50 and keeping ρD=0.95 in the condition where ρC≥0.25, it is observed that AdjAUC values are significantly greater than AUC values. In [ρC=ρD=0]; [ρC=0, ρD=0.95] and [ρC=0.95, ρD=0] combinations, AdjAUC and AUC have similar values. It is observed that, in general, the value of μYD has an effect on AUC. In μYD=0.50 and μYD=1.00 values, it is observed that both AdjAUC and AUC have low values and AdjAUC and AUC values cannot make a good distinction for disease-healthy events.

Table 2. The results of simulated data according to the conditions of $n_D=n_C=250$

μYD	ρC	ρD	AUC	SEMAUC	AdjAUC	SEMAdjAUC	p
1.50	0.25	0.25	0.811	0.015	0.819	0.015	0.627
	0.50	0.50	0.759	0.019	0.791	0.020	0.110
	0.75	0.75	0.702	0.025	0.787	0.031	0.002
	0.95	0.95	0.653	0.030	0.889	0.048	<0.001
	0.95	0.25	0.651	0.031	0.671	0.035	0.538
	0.25	0.95	0.811	0.010	0.844	0.011	0.030
1.70	0.25	0.25	0.847	0.014	0.854	0.014	0.589
	0.50	0.50	0.800	0.017	0.834	0.018	0.058
	0.75	0.75	0.748	0.074	0.843	0.026	<0.001
	0.95	0.95	0.703	0.030	0.951	0.027	<0.001
	0.95	0.25	0.702	0.030	0.728	0.032	0.411
	0.25	0.95	0.847	0.014	0.879	0.014	0.025
1.90	0.25	0.25	0.877	0.013	0.885	0.012	0.560
	0.50	0.50	0.838	0.017	0.872	0.016	0.043
	0.75	0.75	0.790	0.022	0.888	0.021	<0.001
	0.95	0.95	0.749	0.028	0.981	0.012	<0.001
	0.95	0.25	0.749	0.028	0.779	0.030	0.313
	0.25	0.95	0.878	0.013	0.909	0.012	0.016
2.00	0.25	0.25	0.891	0.011	0.898	0.011	0.011
	0.50	0.50	0.854	0.016	0.888	0.015	0.030
	0.75	0.75	0.812	0.022	0.908	0.019	<0.001
	0.95	0.95	0.770	0.026	0.989	0.007	<0.001
	0.95	0.25	0.770	0.027	0.801	0.029	0.278
	0.25	0.95	0.891	0.012	0.921	0.010	0.009

As seen in table 3, it is observed that AdjAUC values are significantly greater than the AUC values for data simulated in nD=nC=100 conditions and in $\mu_{YD}=1.50$ and $\rho_C=\rho D\geq 0.75$ conditions. Besides this, when keeping $\rho_C=0.95$ constant in

the condition where ρD≥0.50, it is found that AdjAUC values are significantly greater than AUC values. In μYD≥1.60 conditions, where ρC=ρD≥0.50, it is found that AdjAUC values are significantly greater than AUC values. Besides this, when keeping ρC=0.95 constant in the condition where ρD≥0.75 and keeping ρD=0.95 constant in the condition where ρC≥0.50, it is found that AdjAUC values are significantly greater than AUC values. In [ρC=ρD=0]; [ρC=0, ρD=0.95] and [ρC=0.95, ρD=0] combinations, AdjAUC and AUC have similar values. It is observed that, in general, the value of μYD has an effect on AUC. When μYD=0.50 and μYD=1.00, it is observed that both AdjAUC and AUC have low values and AdjAUC and AUC values cannot make a good distinction for disease-healthy events.

Table 3. The results of simulated data according to the conditions of n_D=n_C=100

μYD	ρC	ρD	AUC	SEMAUC	AdjAUC	SEMAdjAUC	p
1.50	0.25	0.25	0.812	0.024	0.819	0.024	0.753
	0.50	0.50	0.162	0.032	0.794	0.033	0.327
	0.75	0.75	0.702	0.041	0.786	0.050	0.069
	0.95	0.95	0.651	0.046	0.877	0.076	<0.001
	0.95	0.25	0.649	0.049	0.668	0.055	0.706
	0.25	0.95	0.810	0.024	0.843	0.025	0.196
1.70	0.25	0.25	0.848	0.021	0.855	0.021	0.725
	0.50	0.50	0.800	0.029	0.834	0.029	0.254
	0.75	0.75	0.747	0.039	0.841	0.042	0.022
	0.95	0.95	0.700	0.048	0.939	0.050	<0.001
	0.95	0.25	0.700	0.048	0.726	0.052	0.615
	0.25	0.95	0.846	0.022	0.879	0.021	0.144
1.90	0.25	0.25	0.877	0.020	0.885	0.020	0.716
	0.50	0.50	0.837	0.026	0.871	0.025	0.188
	0.75	0.75	0.791	0.036	0.888	0.034	0.006
	0.95	0.95	0.745	0.045	0.975	0.026	<0.001
	0.95	0.25	0.748	0.047	0.777	0.050	0.050
	0.25	0.95	0.878	0.020	0.908	0.019	0.127
2.00	0.25	0.25	0.891	0.019	0.898	0.019	0.705
	0.50	0.50	0.855	0.024	0.889	0.023	0.160
	0.75	0.75	0.809	0.034	0.905	0.030	0.003
	0.95	0.95	0.767	0.044	0.984	0.019	<0.001
	0.95	0.25	0.768	0.044	0.798	0.046	0.504
	0.25	0.95	0.890	0.019	0.920	0.017	0.111

As seen in table 4, it is observed that AdjAUC values are significantly greater than the AUC values for data simulated in nD=nC=50 conditions and in μYD=1.50 and ρC=ρD≥0.95 conditions. In the conditions where 1.60≤μYD≤1.80 and ρC=ρD≥0.95, it is observed that AdjAUC values are significantly greater than AUC values. Besides this, when keeping ρC=0.95 constant in the condition where ρD≥0.75, it is found that AdjAUC values are significantly greater than AUC values. In μYD=1.90 condition, where ρC=ρD≥0.95, it is found that AdjAUC values are significantly greater than AUC values. Besides this, when keeping ρC=0.95 constant in the condition where ρD≥0.5, it is found that AdjAUC values are significantly greater than AUC values. In μYD=2.00 condition, where ρC=ρD≥0.75, it is found that AdjAUC values are significantly greater than AUC values. Besides this, when keeping ρC=0.95 constant in the condition where ρD≥0.75, and keeping ρD=0.95 constant in the condition where ρC≥0.5, it is found that AdjAUC values are significantly greater than AUC values. In [ρC=ρD=0], [ρC=0, ρD=0.95] and [ρC=0.95, ρD=0] combinations, AdjAUC and AUC have similar values. It is observed that, in general, the value of μYD has an effect on AUC. When μYD=0.50 and μYD=1.00, it is observed that both AdjAUC and AUC have low values and AdjAUC and AUC values cannot make a good distinction for disease-healthy events.

Table 4. The results of simulated data according to the conditions of n_D=n_C=50

μYD	ρC	ρD	AUC	SEMAUC	AdjAUC	SEMAdjAUC	p
1.50	0.25	0.25	0.812	0.034	0.819	0.034	0.823
	0.50	0.50	0.757	0.048	0.788	0.050	0.524
	0.75	0.75	0.700	0.059	0.782	0.072	0.217
	0.95	0.95	0.648	0.068	0.856	0.122	0.029
	0.95	0.25	0.650	0.065	0.670	0.073	0.777
	0.25	0.95	0.810	0.035	0.842	0.035	0.360
1.70	0.25	0.25	0.845	0.031	0.853	0.031	0.808
	0.50	0.50	0.799	0.043	0.833	0.043	0.441
	0.75	0.75	0.744	0.057	0.835	0.064	0.138
	0.95	0.95	0.705	0.067	0.932	0.077	0.002
	0.95	0.25	0.696	0.069	0.721	0.076	0.738
	0.25	0.95	0.846	0.034	0.878	0.033	0.345
1.90	0.25	0.25	0.877	0.028	0.844	0.028	0.793
	0.50	0.50	0.839	0.036	0.873	0.035	0.349
	0.75	0.75	0.787	0.051	0.881	0.049	0.062
	0.95	0.95	0.752	0.064	0.968	0.050	<0.001
	0.95	0.25	0.750	0.064	0.779	0.068	0.669
	0.25	0.95	0.876	0.028	0.907	0.026	0.274
2.00	0.25	0.25	0.891	0.027	0.898	0.027	0.794
	0.50	0.50	0.853	0.036	0.886	0.034	0.343
	0.75	0.75	0.808	0.048	0.902	0.043	0.042
	0.95	0.95	0.769	0.064	0.978	0.037	<0.001
	0.95	0.25	0.770	0.064	0.799	0.067	0.656
	0.25	0.95	0.891	0.027	0.920	0.024	0.258

As seen in table 5, it is observed that AdjAUC values are significantly greater than the AUC values for data simulated in nD=nC=25 conditions and in ρC≥0.75 and ρD≥0.95 conditions. No significant difference could be found between AdjAUC and AUC values in any of other ρC,ρD combinations. It is found that the values of AdjAUC and AUC are similar. It is observed that, in general, the value of μYD has an effect on AUC. When μYD=0.50, it is observed that both AdjAUC and AUC have low values. When μYD=1.00, it observed that it has a distinction greater than of μYD=0.50 value, but still has not effective distinction. When μYD has a value greater than or equal to 1.50, the area under the curve gets larger and for μYD=2.00, it reaches the highest level. If covariate parameters increase as biomarker parameters, the distinction of disease and healthy individuals increases significantly. It is observed that the n=25 value is insufficient AdjAUC calculations except ρC≥0.75 and ρD≥0.95 when considering its sample size.

For all combinations, it is seen that ρ=0.10 value is not sufficient to observe the covariate effect in both disease and healthy group.

Table 5. The results of simulated data according to the conditions of $n_D=n_C=25$

μ_{YD}	ρ_C	ρ_D	AUC	SEM_{AUC}	AdjAUC	SEM_{AdjAUC}	p
1.50	0.25	0.25	0.812	0.049	0.819	0.049	0.881
	0.50	0.50	0.760	0.060	0.790	0.065	0.631
	0.75	0.75	0.698	0.082	0775.	0.100	0.401
	0.95	0.95	0.646	0.098	0.825	0.184	0.208
	0.95	0.25	0.645	0.099	0.663	0.110	0.864
	0.25	0.95	0.807	0.051	0.838	0.052	0.544
1.70	0.25	0.25	0.848	0.045	0.856	0.045	0.870
	0.50	0.50	0.799	0.059	0.832	0.059	0.584
	0.75	0.75	0.747	0.080	0.832	0.089	0.316
	0.95	0.95	0.698	0.096	0.899	0.137	0.085
	0.95	0.25	0.700	0.098	0.723	0.107	0.821
	0.25	0.95	0.846	0.047	0.877	0.045	0.503
1.90	0.25	0.25	0.875	0.041	0.882	0.040	0.860
	0.50	0.50	0.836	0.054	0.869	0.051	0.539
	0.75	0.75	0.789	0.073	0.878	0.073	0.225
	0.95	0.95	0.742	0.095	0.943	0.101	0.042
	0.95	0.25	0.744	0.094	0.771	0.100	0.786
	0.25	0.95	0.874	0.040	0.904	0.037	0.432
2.00	0.25	0.25	0.889	0.038	0.896	0.037	0.852
	0.50	0.50	0.851	0.052	0.883	0.050	0.529
	0.75	0.75	0.805	0.071	0.893	0.068	0.203
	0.95	0.95	0.766	0.088	0.962	0.078	0.019
	0.95	0.25	0.764	0.086	0.792	0.090	0.752
	0.25	0.95	0.889	0.038	0.918	0.034	0.428

According to simulation studies done with nD=nC=10, 0.50≤μYD≤2.00, and μZD=1.00 conditions, it is found that AdjAUC and AUC values are almost same in all the ρC,ρ D combinations. It is observed that the n=10 value is insufficient AdjAUC calculations when considering its sample size.

Results of Adjusted ROC (AdjROC) Application with PCA Data Descriptive statistics about 125 patients who contacted Gaziosmanpasa University Medical Faculty Urology Clinic at the years of 2005 to 2007, whose PSA values were measured, prostate biopsies were done and pathology results were recorded with pre-diagnosis of prostate cancer, are as table 6.

Table 6. Descriptive Statistics of Age and PSA

Variables	Status of Prostate CA						t	p
	Prostate CA			Non-Prostate CA				
	n	Mean	SD	n	Mean	SD		
PSA	25	33.69	33.16	100	11.20	9.19	3.359	0.002
Log-PSA	25	1.34	0.40	100	0.95	0.32	-5.249	<0.001
Age	25	70.96	5.10	100	66.96	7.24	3.197	0.003

A weak correlation (r=0.288, p=0.163 and r=-0.072, p=0.476 respectively) was found between PCA values that has applied logarithmic conversion with age in group with prostate cancer and group without prostate cancer.

After adopting program that has been prepared for simulation for a single data set and applying it in prostate cancer data set that has been given descriptive statistics above; AUC was found as 0.7796 in ROC analysis that is not adjusted by age. In ROC analysis adjusted by age, AdjAUC was found as 0.9995. When age variable which has shown to have a relation with PSA about prostate cancer diagnosis in literature was included into the model as a covariate, it is found that adjusted AUC is bigger than the unadjusted AUC.

Discussion

ROC analysis has been commonly used to distinguish rationally between healthy and disease people using biomarkers since 1990s.

Since almost all of the biological characteristics of the individuals are interrelated, it is necessary to know how to deal with biomarker data alone from laboratory tests or in combination with covariates and how to best put laboratory test to use for diagnostic purposes. For this purpose, adjusted ROC analysis approaches based on the covariate which also take into account the covariates according to proper theoretical constructions were developed.

As seen in simulation studies, both AdjAUC and AUC had low values, and there was not significant difference between the two AUC values when sample sizes of disease and healthy group is 10≤nD=nC≤1000, healthy group parameters for the covariate and biomarker are μYC=1.00, σYC=σZC=1.00, and the ones of disease groups are 0.5≤μYD≤1.00, σYD=σZD=1.00. This fact means that if the biomarker cannot distinguish between disease and healthy individuals, area under the ROC curve is low. For the situations in which biomarker cannot distinguish effectively, covariate adjustment is necessary. In other words, if AUC≤0.85, covariate is needed depending upon the decreases in AUC. If there is a correlation of ρD≥0.50 between the covariate and biomarker, the covariate makes a significant contribution the distinguishing power of biomarker.

For the conditions in which sample sizes of disease and healthy group are 10≤nD=nC≤1000, healthy group parameters for covariate and biomarker are μYC=1.00, σYC=σZC=1.00, and those of disease group are μYD≥1.50, σYD=σZD=1.00, area under the ROC curve is remarkably high. When the biomarker values in disease and healthy group are clearly different (AUC≥0.85) and ρC=ρD≥0.50, no contribution of covariate based adjustment in ROC analysis occurs. If the biomarker distinguishes the disease and healthy individuals, correlation between covariate and biomarker in disease group is very low ($\rho C \rightarrow 0$) and the one in disease group is sufficient (ρD≥0.50),

and at the same time, distinguishing power of biomarker is moderate (AUC≤0.75) without the covariate effect, addition of the covariate into the model seems beneficial.

Janes et al. conducted simulation studies in which they conducted mean increases based on data constructs with standard normal distribution, and reported that mean increases significantly increased AUC value provided that standard deviation values remain constant.[3] It seemed that when it joined to this combination of variables, covariate made significant contributions to AUC value. The results of Janes et al. are parallel to ours.[3] However, it can be seen that they did not change the values of control, covariate and biomarker. In our study, on the other hand, we studied the effects of sample size, mean and correlation coefficient on adjusted ROC analyses, also studying permutational combinations. When the sample size was over 50 (n>50), mean increased and correlation coefficient got larger in these mixed models, AdjAUC values had a tendency toward significantly increased distinguishing among values.

Although the simulation efforts in our study are very rare in literature, the results obtained are similar to those of Janes et al. [3]. In Janes *et al.*'s study, ρC =0.10 and ρD=0.90, and other conditions were similar to the ones in our study. In our study, all other correlation combinations were also tried and more detailed results were obtained.

When the correlations of covariate and biomarker in disease and healthy group is very small (ρC±0.05 and ρD±0.05) and close to each other (ρC≈ ρD), covariate has no contribution to distinguishing between disease and healthy.

When 1000≥nD=nC≥50, expected differences between AUC and AdjAUC values becomes clear. It seems that algorithm prepared works well when nD=nC≥50. When the sample size is insufficient (nD=nC<50), no significant difference is found between AUC and AdjAUC. When nD=nC<50, it seems that AdjROC method is not preferable. This result is contradictory to that of study by Janes et al. , who showed that performance of small sample sizes were satisfactory.[2]

Based on the simulation study in which the sample sizes were different (nD≠nC) and ratio of healthy group to disease group (nD:nC) was 1:2, it was seen that results which were not different from the situation in which sample size were equal (nD=nC) were not different. Thus, sample size did not have any further contribution to AUC.

In the treatment in which age variable, demonstrated to have a relationship with PSA in prostate cancer diagnosis, was included in the model as covariate, there was significant difference between AdjAUC and AUC, and age variable appeared to increase the area under ROC curve (AdjAUC=0.9995, AUC=0.7796). This finding is in accordance with the ones from the simulation study. PSA has a moderate level distinguishing power and has a small correlation with age. This is parallel to the results of the simulation studies, and resulted in a 22% difference between AdjAUC and AUC. In most of the age-adjusted PSA studies, young males are also included. This higher age average in the present study did not prevent revealing of the distinguishing characteristic of age. Our results are parallel to the ones from the studies in literature. 2, 8, 20

Conclusion

In order for AdjROC model to be practical, the distributions of covariate and biomarker need to be binormal distributions. If a biomarker can distinguish well between healthy and disease individuals (AUC≥0.85), there are no need for an additional variable. However, if a biomarker cannot distinguish them (AUC<0.85), there may be a need for covariate depending upon the decrease in AUC.

When the correlations of covariate and biomarker in disease and healthy groups are very small (ρC±0.05 and ρD±0.05) and similar (ρC≈ ρD), there is no contribution of covariate to the model.

When the biomarker distinguishes the disease and healthy individuals well, the correlation between covariate and biomarker in healthy group is very small (ρC→0), the one in disease group is sufficient (ρD≥0.50), and at the same time, distinguishing power of ROC without the covariate effect is moderate (AUC≤0.75), covariate appears to have a positive contribution to the model.

When the distinguishing power of PSA, a prostate cancer biomarker, is evaluated using ROC analysis, AUC adjusted based on age has a higher value compared to unadjusted AUC.

In the model employed, contribution of a single covariate to the diagnostic power of the biomarker was evaluated. It could be beneficial to investigate what the results might be using a different approach in which two or more covariates are employed.

References

1. Janes H. Adjusting for Covariate Effects in Biomarker Studies Using The Subject-Specific Threshold ROC Curve. University Of Washington 2005; Ph.D. Thesis:179.
2. Janes H, Pepe MS. Adjusting for covariate effects on classification accuracy using the covariate-adjusted receiver operating characteristic curve. Biometrika. 2009;96(2):371-382.
3. Janes H, Pepe MS. Matching in studies of classification accuracy: implications for analysis, efficiency, and assessment of incremental value. Biometrics. 2008;64(1):1-9.
4. Ozdamar K. SPSS ile Biyoistatistik. Eskişehir, Turkey: Kaan Kitabevi; 2003.
5. Alonzo TA, Pepe MS. Distribution-free ROC analysis using binary regression techniques. Biostatistics. 2002;3(3):421-432.
6. Cai TX, Pepe MS. Semiparametric receiver operating characteristic analysis to evaluate biomarkers for disease. Journal of the American Statistical Association. 2002;97(460):1099-1107.
7. Faraggi D. Adjusting receiver operating characteristic curves and related indices for covariates. J Roy Stat Soc D-Sta. 2003;52:179-192.
8. Ghosh D, Chinnaiyan AM. Covariate adjustment in the analysis of microarray data from clinical studies. Funct Integr Genomics. 2005;5(1):18-27.
9. 9. Janes H, Pepe MS. Adjusting for covariates in studies of diagnostic, screening, or prognostic markers: an old concept in a new setting. Am J Epidemiol. 2008;168(1):89-97.
10. Janes H, Longton G, Pepe M. Accommodating Covariates in ROC Analysis. Stata J. 2009;9(1):17-39.
11. Le CT. Evaluation of confounding effects in ROC studies. Biometrics. 1997;53(3):998-1007.
12. Pepe MS. Three approaches to regression analysis of receiver operating characteristic curves for continuous test results. Biometrics. 1998;54(1):124-135.
13. Pepe MS. An interpretation for the ROC curve and inference using GLM procedures. Biometrics. 2000;56(2):352-359.
14. Pepe MS. The Statistical Evaluation of Medical Tests for Classification and Prediction. New York: Oxford University Press; 2003.
15. Schisterman EF, Faraggi D, Reiser B. Adjusting the generalized ROC curve for covariates. Stat Med. Nov 15 2004;23(21):3319-3331.
16. Toledano AY, Gatsonis C. Ordinal regression methodology for ROC curves derived from correlated data. Stat Med. 1996;15(16):1807-1826.
17. Tosteson AN, Begg CB. A general regression methodology for ROC curve estimation. Med Decis Making. 1988;8(3):204-215.
18. Cai T. Semi-parametric ROC regression analysis with placement values. Biostatistics. 2004;5(1):45-60.
19. Cai T, Moskowitz CS. Semi-parametric estimation of the binormal ROC curve for a continuous diagnostic test. Biostatistics. 2004;5(4):573-586.
20. Punglia RS, D'Amico AV, Catalona WJ, Roehl KA, Kuntz KM. Effect of verification bias on screening for prostate cancer by measurement of prostate-specific antigen. N Engl J Med. 2003;349(4):335-342.
21. Zhang Z. A Linear Regression Framework for Receiver Operating Characteristic(ROC) Curve Analysis. UW Biostatistics Working Paper Series, Working Paper:253. 2005.
22. Schisterman EF, Reiser B, Faraggi D. ROC analysis for markers with mass at zero. Stat Med. 2006;25(4):623-638.
23. Monson RM. Occupational Epidemiology. Boca Raton, Florida: CRC Press; 1980.