

(نامه به سردبیر)

## کاربرد و تفسیر منحنی خاصه محرکه گیرنده در پژوهش های پزشکی

مجتبی سپندی<sup>۱</sup>، حمید حیدریان میری<sup>۱</sup>، عبدالرضا رجایی فرد<sup>۲\*</sup>

۱- دانشجوی دکترای تخصصی اپیدمیولوژی، گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شیراز، شیراز، ایران.

۲- استاد، گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شیراز، شیراز، ایران.

### چکیده

منحنی مشخصه ی عملکرد (ROC (Receiver operating characteristic که با قرار دادن حساسیت تست تشخیصی در محور  $y$  و  $FPR$  یا میزان مثبت کاذب (۱-ویژگی) در محور  $x$ ها تهیه می شود، روشی مفید برای ارزیابی کارایی تست تشخیصی است. هدف این مقاله، معرفی تحلیل ROC بون استفاده از فرمول های ریاضی و آماری است. در این مقاله همچنین مفاهیم مهم مورد نیاز برای استفاده و تفسیر صحیح این نوع تحلیل ها که عبارتند از تست های تشخیصی کمی و کیفی، حساسیت، ویژگی، شاخص  $AUC$  (Area under the curve) و فاصله اطمینان آن مورد بحث قرار می گیرند. در پایان به اختصار به نرم افزارهایی که تحلیل های ROC را انجام می دهند، اشاره شده است.

واژگان کلیدی: منحنی خاصه محرکه گیرنده، ROC

\*نویسنده مسئول: عبدالرضا رجایی فرد، گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شیراز، شیراز، ایران  
پست الکترونیک:  
msepanidi@gmail.com

تاریخ پذیرش: ۸۹/۸/۳

تاریخ دریافت: ۸۹/۵/۲۵

### مقدمه

یک بیماری واحد و تعیین نقطه ی برش مناسب برای تست هایی که نتیجه ی آن ها در قالب یک عدد بیان می شود (تستهای کمی). این مقاله با ارایه ی مثال هایی کاربردی، مفهوم و کاربردهای منحنی های ROC را به زبانی ساده ارائه می کند.

منحنی های مشخصه ی عملکرد ROC، اولین بار برای تشخیص سیگنال های رادیویی بکارگرفته شدند. بعد ها این منحنی ها در تحقیقات و تصمیم گیری های پزشکی کاربردهای قابل توجهی پیدا کردند (۱و۲). این منحنی ها در سه زمینه کاربرد دارند: تعیین کارایی یک تست مشخص، مقایسه کارایی دو تست مختلف در تشخیص

کارایی یک تست تشخیصی به میزان توافق نتایج آن تست با نتایج استاندارد طلایی بستگی دارد. هرچه مجموع خانه های صحیح (TP و TN) به سمت ۱۰۰٪ و مجموع خانه های کاذب (FP و FN) به سمت صفر میل کند، تست کاربری بیشتری دارد. بنابراین می توان نتیجه گرفت حساسیت و ویژگی دو شاخص مهم برای تعیین کارایی تست تشخیصی شماره می آیند (۳) و برای تعیین کارایی تست، بهتر است این دو شاخص را همزمان مد نظر قرار داد. در مورد تست های کیفی دو حالت به راحتی با محاسبه ی حساسیت و ویژگی می توان در باره ی کارایی تست اظهار نظر نمود، اما وقتی که تست ها، کیفی چند حالت و یا کمی باشند آنگاه باید به ازاء تعداد حالات ممکن، نتیجه ی تست حساسیت و ویژگی محاسبه شود. این جاست که منحنی های ROC می توانند بسیار مفید واقع شوند.

فرض کنید ۷۰ بیمار رادیو گرافی سینه انجام داده اند و مشخص شده است که مشکلات ریوی در آنها وجود دارد، بر اساس نظر رادیولوژیست، آنها به ۵ گروه، خوش خیم، خوش خیم احتمال قوی، خوش خیم با احتمال ضعیف، بدخیم با احتمال قوی، بدخیم تقسیم شده اند. برای تعیین خوش خیم یا بدخیم بودن مشکلات ریوی، بیوپسی انجام می شود (جدول ۲). در این مثال می توان ۴ نقطه برش انتخاب کرد (بزرگتر یا مساوی ۲، بزرگتر یا مساوی ۳، بزرگتر یا مساوی ۴، بزرگتر یا مساوی ۵) و برای هر یک از چهار نقطه یک حساسیت و یک ویژگی محاسبه نمود.

#### نحوه ی رسم و تفسیر منحنی ROC

در منحنی ROC میزان مثبت کاذب (FPR) یا False Positive Rate یا همان (ویژگی -۱) در محور افقی و حساسیت در محور عمودی قرار می گیرد. حال می توان با استفاده از داده های جدول ۲ منحنی ROC را رسم نمود (شکل ۱ و ۲). چنین منحنی هایی را منحنی ROC

#### چند مفهوم پایه ای:

آزمایش ها بر اساس اینکه چگونگی گزارش نتایج آنها به دو گروه کمی و کیفی تقسیم می شوند. نتایج آزمایش های کیفی با مقادیر عددی مشخص نمی شوند. از آن جمله می توان به آزمایش هایی مثل وجود یا عدم وجود انٹی ژن هپاتیت B در سرم اشاره کرد. اگرچه نتایج آزمایش های کمی در قالب عدد بیان می شود (مثلا  $145 \pm 8 \text{ mg/dl}$  = کلسترول) اما می توان آن ها را به صورت کیفی نیز ارایه کرد (مثل طبیعی و غیر طبیعی یا مثبت و منفی). کارایی یک آزمایش کمی وقتی به کیفی تبدیل می شود به این بستگی دارد که در چه نقطه ای، آزمایش غیر طبیعی یا مثبت در نظر گرفته می شود (۳). برای تعیین کارایی یک آزمایش به یک ملاک بیماری مثل کم کاری مادرزادی تیروئید آزمایش کنیم. تستی را در نظر بگیرید که به سهولت و سرعت قابل انجام است و وجود یا عدم وجود بیماری را مشخص می کند؟ مثلا در انجام تست TSH در نمونه های به دست آمده از پاشنه پای نوزادان، با فرض مشخص بودن تعداد نوزادان چنین عمل می کنیم: اولاً تست TSH را انجام داده و نتایج را ثبت می کنیم، ثانياً با روشهای کاملاً مطمئن و دقیق که احتمالاً گران تر هم هستند (مثلاً TSH و T4) سرمی و همچنین با کمک پزشک متخصص و بدون استفاده از تست TSH پاشنه ی پا، وجود یا عدم وجود کم کاری مادرزادی تیروئید نوزادان را بررسی کرده و دوباره نتیجه ها را ثبت می کنیم (جدول ۱).

**حساسیت (Sensitivity):** عبارتست از توانایی تست برای تشخیص درست نوزادان بیمار،  $+FN = TP/TP$  حساسیت

**ویژگی (Specificity):** عبارتست از توانایی تست برای تشخیص درست نوزادان سالم. در واقع نسبت تشخیص های درست عدم بیماری به کل نوزادانی که فاقد بیماری هستند (۳).  $TN/TN+FP =$  ویژگی

درصد هم احتمال دارد همان فرد را در گروه سالمها قرار دهد. **AUC** را هم با روشهای پارامتری (با فرض **Binormal** بودن نتایج تست) و هم با روشهای نا پارامتری می توان برآورد کرد(۹). مسئله ی دیگر در باره **AUC** لزوم محاسبه و گزارش فاصله ی اطمینان برای آن است (۱, ۸). از آنجا که **AUC** شاخصی از کارایی کلی تست است. را می توان کارایی دو تست مختلف با مقایسه **AUC** های آن ها مقایسه ی کرد. برای مقایسه دو **AUC** می توان از تست **Z** استفاده کرد(۱۰).

نرم افزارهای **SPSS** ، **SAS** ، **STATA** ، **ROCKIT** ، **PlotROC.xls** و **MedCalc** ، برآورد منحنی های **ROC**، محاسبه آماری **AUC** و خطای معیار آن را انجام می دهند. به عنوان نتیجه از آن جا که برای ارزیابی و مقایسه تست های تشخیصی روش آماری مناسبی به جز استفاده از منحنی های **ROC** وجود ندارد، لذا در این مقاله سعی بر آن بود به سادگی و با اجتناب از مباحث پیچیده ی آماری و البته باحفظ مفاهیم پایه ای، کاربرد های آن در حوزه ی تحقیقات پزشکی برای محققین معرفی شود.

هموار شده (**Fitted or Smooth ROC Curve**) می نامند (۱۰) (شکل ۲). نکته ی جالب توجه و مهم این است که از آنجا که منحنی رسم شده مربوط به تمام نقاط برش ممکن است، می توان بدون توجه به نقطه برش به طور کلی تست تشخیصی را ارزیابی کرد. هرچه منحنی رسم شده به بالا و سمت چپ (حساسیت معادل ۱) نزدیکتر (یا از نیمساز ۴۵ درجه بین دو محور دور تر) باشد کارایی آن بیشتر است، عکس قضیه نیز صادق است. چندین شاخص برای ارزیابی تست در رابطه با منحنی **ROC** وجود دارد که از همه معروفتر سطح زیر منحنی یا **AUC** است(۱۰). درواقع **AUC** را می توان مقدار متوسط حساسیت برای تمام نقاط برش ممکن دانست که عددی بین صفر و یک است. هر قدر مقدار **AUC** به یک نزدیکتر باشد به آن معناست که تست بهتر عمل می کند( و بالعکس). اما حد اقل مقدار **AUC** که در عمل استفاده می شود، ۰/۵ است زیرا استفاده از تستی که **AUC** آن معادل ۰/۵ است، نمی توان در مورد تعیین بیماری افراد نظر قطعی داد(۵, ۷)، به این معنا که استفاده از تست برای تشخیص بیماری در یک فرد معین ۵۰ درصد احتمال دارد فرد در گروه بیماران قرار گیرد و ۵۰

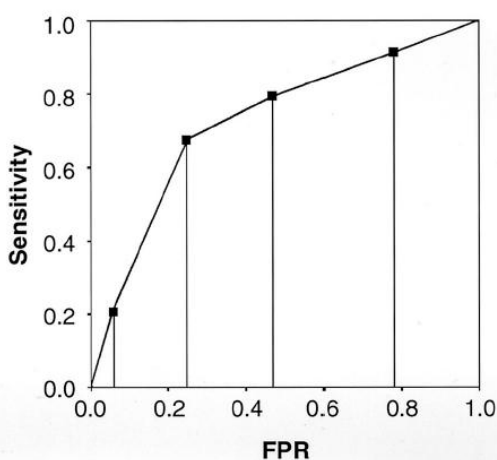
جدول ۱: نتیجه تست **TSH** پاشنه ی پا در مقابل نتیجه تست استاندارد طلائی

		نتیجه ی استاندارد طلائی	
		بیماری وجود ندارد ( _ )	بیماری وجود دارد (+)
نتیجه <b>TSH</b> پاشنه پا	<b>TSH</b> غیر طبیعی (+)	<b>FP</b>	<b>TP</b>
	<b>TSH</b> طبیعی ( _ )	<b>TN</b>	<b>FN</b>

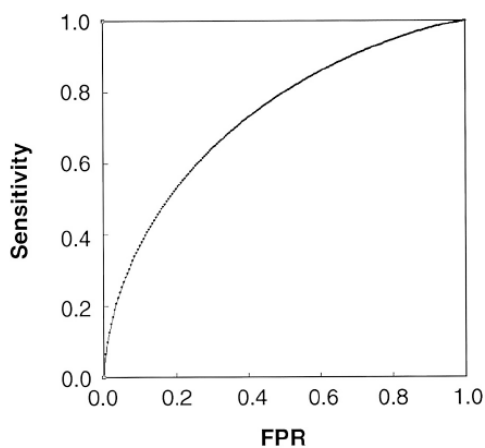
**TP**: مثبت حقیقی یا تشخیص های درست بیماری با استفاده از نتایج تست **TSH**.  
**FP**: مثبت کاذب یا تشخیص های غلط بیماری با استفاده از نتایج تست **TSH** (نوزادان سالمی که بیمار تشخیص دادیم).  
**TN**: منفی حقیقی یا تشخیص های درست فقدان بیماری با استفاده از نتایج تست **TSH**.  
**FN**: منفی کاذب یا تشخیص های غلط فقدان بیماری با استفاده از نتایج تست **TSH** (نوزادان بیماری که سالم تشخیص دادیم).

جدول ۲: مقایسه ی نتایج تست رادیوگرافی و بیوپسی (استاندارد طلایی) و حساسیت و ویژگی تست در نقاط برش مختلف در ۷۰ بیمار .

نتیجه ی بیوپسی	تفسیر رادیولوژی (نقاط برش)				
	۱	۲	۳	۴	۵
خوش خیم	۳	۴	۴	۱۶	۷
بدخیم	۸	۱۱	۸	۷	۲
حساسیت	.....	۰/۹۱	۰/۷۹	۰/۶۷	۰/۲۰
ویژگی-1		۰/۷۷	۰/۴۷	۰/۲۵	۰/۰۵



شکل ۱: نمودار ROC که با وصل کردن نقاط حساسیت و میزان مثبت کاذب (FPR) به ازاء هر نقطه برش به یکدیگر رسم شده است.



شکل ۲: منحنی هموار شده ROC

## References

## منابع

- 1- Campbell G, Levy D, Bailey JJ. Bootstrap comparison of fuzzy ROC curves for ECG-LVH algorithms using data from the Framingham Heart Study. *J Electrocardiol.* 1990; 23 Suppl: 132-7.
- 2- Egan J. *Signal detection theory and ROC analysis*, Series in Cognition and Perception. Academic Press, 1975.
- 3- Gordis L. *Epidemiology*, 4th Edition. Philadelphia, WB Sanders, 2008
- 4- Hanley JA. The robustness of the "binormal" assumptions used in fitting ROC curves. *Med Decis Making.* 1988 Jul-Sep;8(3):197-203
- 5- Metz CE. Basic principles of ROC analysis. *Semin.Nucl.Med.* 1978; 8:283-98.
- 6- Metz CE, Herman BA, Shen JH. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. *Stat Med.* 1998 May 15; 17(9):1033-53.
- 7- Obuchowski NA. Receiver operating characteristic curves and their use in radiology. *Radiology.* 2003Oct;229(1):3-8
- 8- Reiser B, Faraggi D. Confidence intervals for the generalized ROC criterion. *Biometrics.* 1997 Jun; 53(2):644-52.
- 9- Thompson ML, Zucchini W. On the statistical analysis of ROC curves. *Stat.Med.* 1989; 8:1277-90.
- 10- Zhou X, Obuchowski N, McClish D. *Statistical methods in diagnostic medicine*. New York: John Wiley & Sons, 2002.

## Receiver operating characteristic curve's application and interpretation in medical research

Mojtaba Sepandi<sup>1</sup>, Hamid Heidarian Miri<sup>1</sup>, Abdolreza Rajaeefard<sup>2\*</sup>

1- PhD student of Epidemiology, Department of Epidemiology, Faculty of Health, Shiraz University of Medical Sciences, Shiraz, Iran.

2- Professor of Biostatistics, Department of Epidemiology, Faculty of Health, Shiraz University of Medical Sciences, Shiraz, Iran.

---

### Abstract

Receiver operating characteristic(ROC), a sensitive diagnostic test that Plots ssensitivity in the y-axis and the rate of false positives in the x-axis, is prepared to be a useful method for evaluating performance of the diagnostic test. The aim of this study was to introduce the ROC analysis in a non-mathematical and non-statistical way. Important concepts that are needed to use and to interpret this type of analysis correctly include quantitative and qualitative diagnostic tests, sensitivity and specificity, area under the curve, and 95 % confidence intervals. Finally, the ROC analysis software has also been discussed in brief

**Key words:** Receiver operating characteristic (ROC) curve, receiver operating characteristic curve

---

\*Corresponding author: Abdolreza Rajaeefard, PhD, Professor of Biostatistics, Department of Epidemiology, Shiraz University of Medical Sciences, Shiraz, Iran.

Received: August 16, 2010

Accepted: October 25, 2010