



Klinik Karar Vermede ROC Analizi

Selim Kılıç¹

ÖZET:

Klinik karar vermede ROC analizi

ROC eğrisi duyarlılık ve seçicilik arasındaki ilişkinin grafiksel bir gösterimidir. Eğriler yardımı ile bir tanı testi için en iyi eşik değeri saptanarak en uygun modele karar verebiliriz. Ayrıca farklı tanı testlerinin doğru klinik tanı koymadaki başarısının karşılaştırılmasına olanak sağlarlar. ROC analizi; tanı sürecinin uzun zaman alacağı, maliyetin yüksek, özel yöntem-ekipman ve nitelikli insan gücüne gereksinim duyulacağı durumlarda; kısa zamanda, düşük maliyetle, kolay elde edilebilen belirteçler için uygun kesim noktalarını belirleyerek klinik karar verme sürecine önemli katkı sağlayacak bir analiz yöntemidir.

Anahtar sözcükler: doğru pozitif, yanlış pozitif, eğri, eşik değeri

ABSTRACT:

ROC analysis in clinical decision making

ROC curve is a graphic presentation of the relationship between both sensitivity and specificity. By the help of curves we can decide the optimal model through determining the best threshold for a diagnostic test. They also provide comparison the success of different tests in correct clinical diagnosis. ROC analysis is an analysis method that will contribute to the process of clinical decision-making when the diagnosis process will take a long time, the cost will be high, special method-equipment and qualified human resources will be needed by determining appropriate cut-off values for indicators that will be determined in short-time, low-cost, and easily obtainable.

Key words: true positive, false positive, curve, threshold value

Journal of Mood Disorders 2013;3(3):135-40

¹MD, Gülhane Askeri Tıp Fakültesi,
Ankara-Türkiye

Yazışma Adresi / Address reprint requests to:
Selim Kılıç, Gülhane Askeri Tıp Fakültesi,
Ankara-Türkiye

Elektronik posta adresi / E-mail address:
drselimkili@gmail.com

Kabul tarihi / Date of acceptance:
15 Ağustos 2013 / August 15, 2013

Bağıntı beyanı:

S.K.: Yazar bu makale ile ilgili olarak herhangi bir çıkar çatışması bildirmemiştir.

Declaration of interest:

S.K.: The author declare that they have no conflict of interests regarding the content of this article.

GİRİŞ

İlk ROC eğrisi 2. Dünya Savaşı'nda "radar sinyalleri"nin analizi için kullanılmıştır. Düşman uçaklarını, radar sinyallerini kullanarak daha doğru bir şekilde saptamak amacıyla araştırmalara başlanmıştır (1,2). ROC eğrileri tıpta 1960'larda kullanılmaya başlanmıştır. ROC eğrileri tanı testlerinin değerlendirilmesinde özellikle radyolojide ve psikolojide sıkça kullanılmaktadır (3-7).

ROC eğrileri sonuç değişkeninin özellikle iki olasılıklı (depresyon var-yok, remisyon var-yok, nüks var-yok,...vb) olduğu, buna karşılık karar vermede kullanılacak değişkenin sürekli olduğu durumlarda (kortizol, glisemi düzeyi gibi) kullanılırlar. ROC eğrileri bu sürekli değişken için olası tüm kesim noktalarını gösterir ve her kesim noktasında değişik sonuçların - doğru pozitif (DP), doğru negatif (DN), yanlış pozitif (YP) ve yanlış negatif (YN)- sıklığı

hakkında kestirimler yapılmasına olanak sağlarlar. Verilen bir test için en iyi kesim noktasının saptanmasında, doğru ve yanlış kararların yararlarının hesaplanmasında kullanılırlar. Ayrıca farklı tanı testlerinin doğru tanı koymadaki başarısının karşılaştırılmasına olanak sağlarlar (2,8-11).

	Hastalık (+)	Hastalık (-)
Test (+)	Doğru pozitif (a)	Yanlış pozitif (c)
Test (-)	Yanlış negatif (b)	Doğru negatif (d)
	n1=(a+b)	n2=(c+d)

Doğru Pozitif Orantısı (DPO) = Doğru Pozitif / n1 = duyarlılık (sensitivite)

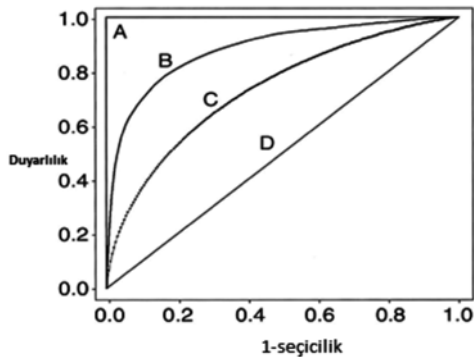
Yanlış Pozitif Orantısı (YPO) = Yanlış Pozitif / n2 = 1- seçicilik (spesifisite)

ROC eğrilerinde; x ekseninde YPO (Yanlış Pozitif Orantısı) yer alırken, y ekseninde DPO (Doğru Pozitif Orantısı) bulunmaktadır. Farklı eşik değerleri için DPO ve YPO değerleri yani duyarlılık ve 1-seçicilik değerleri hesaplanır. DPO ve YPO ikilileri ROC eğrisini oluşturur. ROC Eğrisi (0,0) ile (1,1) arasında artan bir fonksiyondur. ROC analizi sonrası tanı koymadaki başarısı değerlendirilen bir testin yararsız mı yoksa çok iyi bir test mi olduğu kararı verilebilir.

Yararsız Test: Tanı testi hasta olanları ve sağlıklı olanları birbirinden ayıramıyor ise o test yararsız bir testtir ve para ile yazı-tura atmaktan farksız bir olasılığa (%50) sahiptir

Mükemmel Test: Tanı testi hasta olanları ve sağlıklı olanları birbirinden tam anlamıyla ayırabiliyor ise o test mükemmel bir testtir. Bu durum **DPO(c)=1, YPO(c)=0** şeklinde gösterilir. Testlerin birçoğu yararsız test ile mükemmel test arasında bir performansa sahiptir. ROC grafiğinin sol üst köşesine yaklaşıldıkça ((0,1) noktası) testlerin ayırt etme gücü artar, mükemmelliğe yaklaşırlar. Testlerin doğru karar vermede gücünü değerlendirmede kullanılan ölçütlerden biri de "ROC Eğrisi Altında Kalan Alan"dır. Eğri altında kalan alan (area under the curve=AUC) en büyük "1" değerini alabilir. Pratik olarak alabileceği en küçük değer ise "0.50" dir. Hastalarla sağlıklılar tamamen şansa bağlı olarak (örneğin para atışı ile) ayırt edilirse bu durum ortaya çıkar (8,12).

Aşağıda bir grafik üzerinden farklı testlerin karar vermedeki başarısını değerlendirelim.



ROC eğrilerinde daha kuzey batıdaki nokta daha iyidir. İstenilen DPO yüksek iken YPO değerinin düşük olmasıdır. Koordinat olarak (0,1) noktası en iyi sınıflandırmayı gösterir. Bu bilgiler doğrultusunda;

* A testi (0,1) noktası ile mükemmel bir testtir çünkü ROC eğrisi altında kalan alan 1 dir. Zira, x eksen koordinatı (1-seçicilik değerinin karşılığı) olan 0 değeri testin seçiciliğinin %100 olduğunu gösterirken, y eksen koordinat değeri 1 ise testin duyarlılığının %100 olduğunu göstermektedir.

* Buna karşılık D testi hasta olan ve olmayan bireyleri ayırmakta en başarısız testtir. ROC eğrisi altında kalan alanı 0.5 tir.

Kesim noktasına karar verirken değerlendirmeye alınması gereken kriterler

ROC eğrisi yardımıyla en iyi kesim noktasını belirlemede iki farklı yaklaşım kullanılabilir. İlk yaklaşım eğrinin grafiğın sol üst köşesine yani (0,1) koordinatlarına en yakın olduğu noktayı kesim noktası almaktır. İkinci yaklaşımda ise araştırmacı tarafından test için öngörülen bir duyarlılık, seçicilik değeri yoksa yani duyarlılığı en az %95 olmalı, sağlıklıları hiç atlamamalı (seçicilik=%100) gibi bir kriter belirlenmemişse duyarlılık ve seçicilik değerleri toplamının en yüksek bulunduğu nokta en iyi kesim noktası olarak belirlenir (8). Belirlenen noktanın x (1-seçicilik) veya y (duyarlılık) eksenlerinde nerede olması gerektiği kararını verirken birden fazla kriteri dikkate almak gerekir. Bunlardan başlıcaları:

*Hastalığın tedavi maliyeti ve aynı zamanda tedavinin başarısız olma durumundaki maliyetin göz önüne alınması,

*Sonraki aşamalarda tetkiklerin maliyeti,

*Hastaya tedavi sırasında verdiği rahatsızlık veya tedavinin başarısızlığının hastada oluşturacağı etki,

*Tedavi etme veya etmeme durumundaki ölüm riski gibi durumlardır.

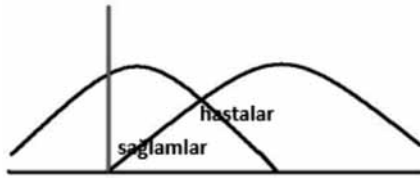
Kesim noktasını alırken yüksek veya düşük bir değeri kabul etmek farklı sonuçlara neden olacaktır. Yüksek değerlerin hastalığı gösterdiği bir durumda taramada kullanılan test için düşük kesim noktası alındığında;

* Hastaların tamamı bulunacak

*Sağlam olanların bir kısmı hasta tanısı alacak (yanlış pozitif)

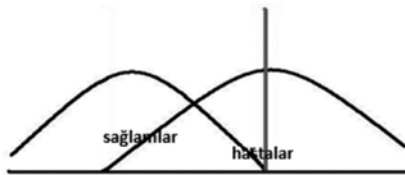
*Tarama testinin **duyarlılığı artacak**

* Buna karşılık ise tarama testinin **seçiciliği azalacaktır.**



Buna karşılık taramada kullanılan test için yüksek kesim noktası alındığında ise;

- *Sağlamların tamamı bulunacak
- *Hastaların bazıları **sağlam** tanısı alacak (yanlış negatif)
- * Tarama testinin **seçiciliği artacak**
- * Buna karşılık ise tarama testinin **duyarlılığı azalacaktır.**



SPSS'de ROC Analizi ve Sonuçların Yorumlanması

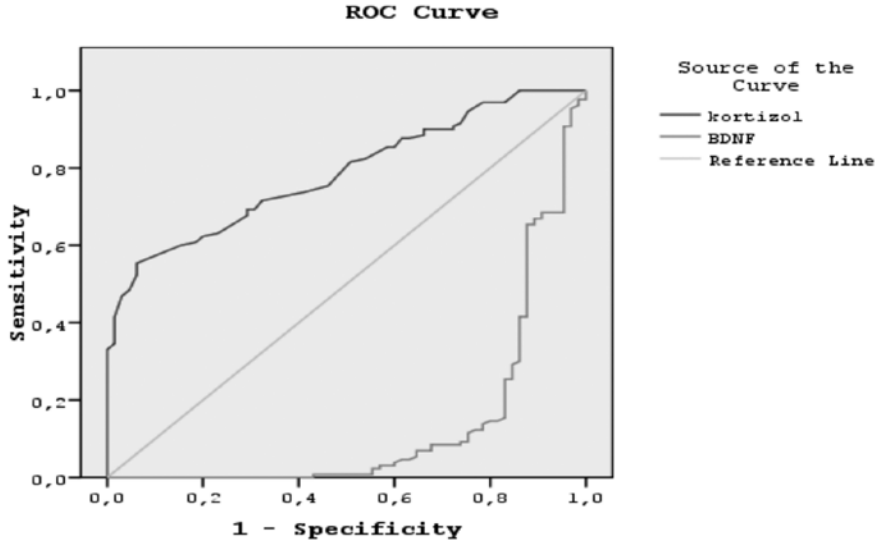
Klinik tanısı depresyon olan ve olmayanlardan oluşan çalışma grubunda tarama testi olarak brain-derived neurotrophic factor (BDNF) ve kortizol değerlerini ROC analizi ile değerlendirelim. SPSS veri tabanı sayfasında “Analyze” seçeneği altındaki ROC Curve seçeneğini işaretliyoruz.

Çıkan pencerede bağımlı değişkenimiz depresyon grubu “state variable” kutucuğuna aktarıyoruz. Depresyon grup değişkeninde depresyonu olanlar “1” olarak kodlandığı için “value of state variable” kutucuğuna 1 yazıyoruz. Tanı koymada başarısını değerlendireceğimiz değişkenler olan BDNF ve kortizolü de “test variable” kutucuğuna aktarıyoruz. Display yazısının altındaki kutucukları da işaretliyoruz ve OK tuşuna basıyoruz.

Çıktı sayfasında grafiğimiz çıkıyor. “With diagonal reference line” seçeneğini işaretlediğimiz için en kötü durum olan AUC=0.50'ye karşılık gelen çizgide grafikte

Name	ROC Curve	Columns	Align	Measure
1	sira_no	8	Right	Nominal
2	hamilton_puan	8	Right	Scale
3	grup	12	Right	Ordinal
4	dogumtarihi	8	Right	Scale
5	cinsiyet	8	Right	Nominal
6	vki	8	Right	Scale
7	HB	8	Right	Scale
8	hb1	8	Right	Scale
9	tsh	8	Right	Scale
10	glikoz	7	Right	Scale
11	madde_kullanimi	7	Right	Scale
12	madde_kullanimi	12	Right	Nominal
13	aile_oykusu	10	Right	Nominal
14	tsh1	9	Right	Scale
15	tsh2	10	Right	Scale
16	kortizol	10	Right	Scale
17	BDNF	8	Right	Scale
18	depresyon_grup	8	Right	Scale
19		10	Right	Scale
20		10	Right	Scale
21		10	Right	Scale
22		10	Right	Scale
23		10	Right	Scale
24		10	Right	Scale
25		10	Right	Scale
26		10	Right	Scale
27		10	Right	Scale
28		10	Right	Scale
29		10	Right	Scale
30		10	Right	Scale
31		10	Right	Scale
32		10	Right	Scale
33		10	Right	Scale
34		10	Right	Scale
35		10	Right	Scale
36		10	Right	Scale
37		10	Right	Scale
38		10	Right	Scale
39		10	Right	Scale
40		10	Right	Scale
41		10	Right	Scale
42		10	Right	Scale
43		10	Right	Scale
44		10	Right	Scale
45		10	Right	Scale
46		10	Right	Scale
47		10	Right	Scale
48		10	Right	Scale
49		10	Right	Scale
50		10	Right	Scale
51		10	Right	Scale
52		10	Right	Scale
53		10	Right	Scale
54		10	Right	Scale
55		10	Right	Scale
56		10	Right	Scale
57		10	Right	Scale
58		10	Right	Scale
59		10	Right	Scale
60		10	Right	Scale
61		10	Right	Scale
62		10	Right	Scale
63		10	Right	Scale
64		10	Right	Scale
65		10	Right	Scale
66		10	Right	Scale
67		10	Right	Scale
68		10	Right	Scale
69		10	Right	Scale
70		10	Right	Scale
71		10	Right	Scale
72		10	Right	Scale
73		10	Right	Scale
74		10	Right	Scale
75		10	Right	Scale
76		10	Right	Scale
77		10	Right	Scale
78		10	Right	Scale
79		10	Right	Scale
80		10	Right	Scale
81		10	Right	Scale
82		10	Right	Scale
83		10	Right	Scale
84		10	Right	Scale
85		10	Right	Scale
86		10	Right	Scale
87		10	Right	Scale
88		10	Right	Scale
89		10	Right	Scale
90		10	Right	Scale
91		10	Right	Scale
92		10	Right	Scale
93		10	Right	Scale
94		10	Right	Scale
95		10	Right	Scale
96		10	Right	Scale
97		10	Right	Scale
98		10	Right	Scale
99		10	Right	Scale
100		10	Right	Scale

Name	Type	Values	Missing	Columns	Align	Measure
1	sira_no	Numeri		8	Right	Nominal
2	hamilton_puan	Numeri		8	Right	Scale
3	grup	Numeri	{0, kontrol}...	12	Right	Ordinal
4	dogumtarihi	Numeri		8	Right	Scale
5	cinsiyet	Numeri	{1, erkek}...	8	Right	Nominal
6	vki	Numeri		8	Right	Scale
7	HB	Numeri		8	Right	Scale
8	hb1	Numeri		8	Right	Scale
9	tsh	Numeri		7	Right	Scale
10	glikoz	Numeri		7	Right	Scale
11	madde_kullanimi	Numeri	{0, yok}...	12	Right	Nominal
12	aile_oykusu	Numeri	{0, yok}...	10	Right	Nominal
13	tsh1	Numeri		9	Right	Scale
14	tsh2	Numeri		10	Right	Scale
15	kortizol	Numeri		8	Right	Scale
16	BDNF	Numeri		8	Right	Scale
17	depresyon_grup	Numeri	{0, yok}...	10	Right	Scale
18			{0, yok}...	17	Right	Scale
19						
20						
21						
22						
23						
24						
25						
26						
27						
28						
29						
30						
31						
32						
33						
34						
35						
36						
37						
38						
39						
40						
41						
42						
43						
44						
45						
46						
47						
48						
49						
50						
51						
52						
53						
54						
55						
56						
57						
58						
59						
60						
61						
62						
63						
64						
65						
66						
67						
68						
69						
70						
71						
72						
73						
74						
75						
76						
77						
78						
79						
80						
81						
82						
83						
84						
85						
86						
87						
88						
89						
90						
91						
92						
93						
94						
95						
96						
97						
98						
99						
100						



Diagonal segments are produced by ties.

Area Under the Curve

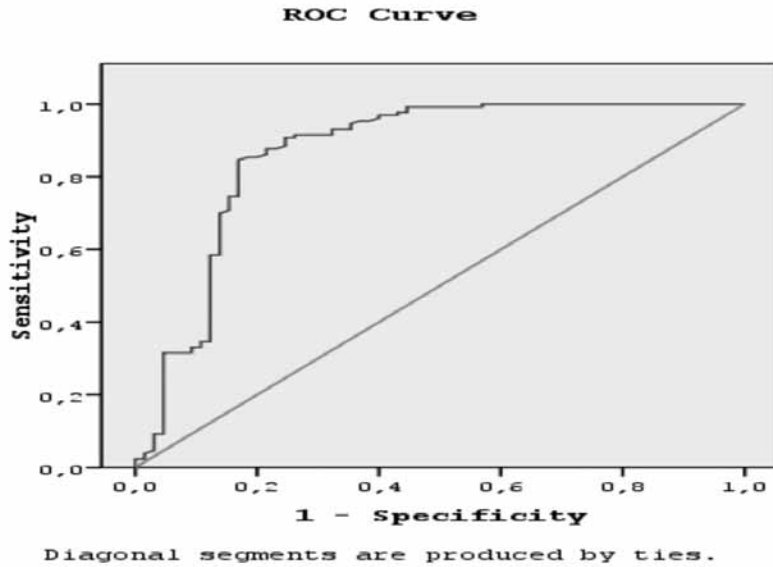
Test Result Variable(s)	Area	Std. Error ^a	Asymptotic Sig. ^b	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
kortizol	,785	,032	,000	,723	,847
BDNF	,134	,033	,000	,069	,199

gözüküyor. Altta kutucukta ise BDNF ve kortizol için AUC değerleri, %95 güven aralıkları ve p değerleri (asympt. sig.) verilmiş. Buna göre kortizol için AUC değeri 0.785 ve $p < 0.001$ yani istatistiksel olarak önemli bulunmuş. AUC değeri için en kötü durumun 0.5 olduğunu belirtmiştik. Peki bu durumda BDNF için bulunan 0.134 değerini nasıl yorumlayacağız?

İstatistik programı yazılımında hastalık kararı vermede yüksek değerler hastalık durumunu gösterir şeklinde

kodlama yapıldığı için bu şekilde bir sonuç ile karşılaştık. Gerçekte ise depresyon grubunda BDNF değerleri depresyon olmayanlardan düşüktür. Yani BDNF'nin düşük olması depresyon olma olasılığını artıran bir durumdur. SPSS programı bize bu konuda değişiklik yapma seçeneğini sunmaktadır. Daha önce açılan pencerede sağ en alta yer alan "options" seçeneğini işaretleyeceğiz ve OK tuşuna basacağız.

Options seçeneğini işaretledikten sonra açılan yeni



Area Under the Curve

Test Result Variable(s): BDNF

Area	Std. Error ^a	Asymptotic Sig. ^b	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.866	.033	.000	.801	.931

sayfada programın daha önce belirlediği “larger test result indicates more positive test” seçeneğini değiştirip “smaller test result indicates more positive test” seçeneğini işaretleyip continue tuşuna basacağız.

Görüldüğü gibi BDNF için olan eğri bu sefer 0.5 referans çizgisinin soluna geçti. Bulduğumuz AUC değeri de bir önceki çıktıda bulduğumuz 0.134 yerine ise $1-0.134=0.866$ olarak hesaplandı. Bu durumda BDNF, AUC değeri 0.794 olan kortizole göre depresyon tanısı koymada daha başarılı bir belirteçtir. SPSS’de ROC analizi

uygulanması ve sonuçların yorumlanmasına ilişkin daha ayrıntılı bilgilere internet ortamından ve basılı kaynaklardan ulaşılabilir (2,13).

Sonuç olarak ROC analizi; tanı sürecinin uzun zaman alacağı, maliyetin yüksek, özel yöntem-ekipman ve nitelikli insan gücüne gereksinim duyulacak durumlarda; kısa zamanda, düşük maliyetle, kolay elde edilebilen belirteçler için uygun kesim noktalarını belirleyerek klinik karar verme sürecine önemli katkı sağlayacak bir analiz yöntemidir.

Kaynaklar:

- Mason SJ, Graham NE. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. Quarterly Journal of the Royal Meteorological Society. 2002;128:2145–66.
- Hayran M, Hayran M. ROC Analizi, Sağlık Araştırmaları İçin Temel İstatistikler. Omega Araştırma Yayınları. Ankara, 2011, 403-416.
- Hanley JA, McNeil BJ. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. Radiology.1982;143: 29–36.
- Obuchowski NA. ROC Analysis. American Journal of Roentgenology. 2005;184:364-72.
- Ratcliff R, McCoon G, Tindall M. Empirical generality of data from recognition memory ROC functions and implications for GMMs. Journal of Experimental Psychology: Learning, Memory, and Cognition.1994;20:763–85.
- Zhang J; Mueller ST. A note on ROC analysis and non-parametric estimate of sensitivity. Psychometrika. 2005;70:203-12.
- Swets JA. Indices of Discrimination or Diagnostic Accuracy Signal Detection Theory and Roc Analysis in Psychology and Diagnostics: Collected Papers. Lawrence Erlbaum Associates. 1996, USA, 59-62.
- Alpar R. Basit Doğrusal Regresyon Çözümlemesi. Spor, Sağlık ve Eğitim Bilimlerinden Örneklerle Uygulamalı İstatistik ve Geçerlik Güvenirlik. Detay Yayıncılık, Ankara, 2010, 338-42.

9. Kirkwood BR, Sterne JAC. Regression Modelling. Medical Statistics. Blackwell Science. 2003. Australia, 432-3.
10. Dawson B, Trapp RG. Statistical Methods for Multiple Variables. Basic & Clinical Biostatistics. Lange Medical Books/McGraw Hill Medical Publishing Division, 2001, USA, 274-76.
11. Pagano M, Gauvreau K. Simple Linear Regression. Principles of Biostatistics . Duxbury Press, 1993, USA, 128-30.
12. Hulley SB, Cummings SR, Browner WS, Grady D, Hearst N, Newman TB. Studies of the Accuracy of Tests. Designing Clinical Research An Epidemiologic Approach. Second Edition. Lippincott Williams&Wilkins, 2001, USA, 181-2.
13. <http://calcnet.mth.cmich.edu/org/spss/StaProcROCCurve.htm>. Erişim tarihi 15.08.2013.