

# Tanı testlerinde sınır değerlerinin belirlenmesi

## The determination of reference values in biomarkers

Hakan ÇAMLICA, Rian DIŞÇI

*İstanbul Üniversitesi Onkoloji Enstitüsü, Kanser Epidemiyolojisi ve Biyoistatistik Bilim Dalı*

Kanserin erken tarama ve tanısına yardımcı ayırmalar kısaca “tanı testleri” olarak adlandırılmaktadır. Bu testlerin iç ve dış geçerlilikleri yüksek olmalıdır. Klinikte rutin kullanılan tanı testlerinin evrensel alt-üst sınır değerleri belirlenmiştir. Fakat pek çok yeni test için bu sınırları belirlemek gerekmektedir. Sınır değerleri parametrik ve non-parametrik istatistik yöntemlerle kolayca saptanabilir. Tıbbi testlerin performanslarının değerlendirilmesinde en önemli ve kullanışlı teknik “Receiver Operating Characteric” (ROC) eğrisi yöntemidir. ROC eğrisi tanı testlerinin güvenilirliği ve genel doğruluğu (*accuracy*) açısından standart yöntem haline gelmiştir.

**Anahtar sözcükler:** Biyometri/yöntemler; duyarlılık; genel doğruluk; sınır değerleri; özgüllük; ROC.

Biomarkers are important molecular signatures of the phenotype of a cell that aid in early cancer detection and risk assessment. The diagnostic attribute of a test, diagnostic sensitivity and diagnostic specificity are important in accessing whether a test is useful. The concept of reference values was introduced to avoid the ambiguities inherent in the term “cut-off values”. The reference interval is easily determined by both parametric and nonparametric statistical methods. An important and useful technique for evaluating the performance of diagnostic medical tests is the receiver operating characteristic (ROC) curve. ROC curves have become the standard for describing and comparing the accuracy of diagnostic tests.

**Key words:** Accuracy; biometry/methods; reference values; ROC; sensitivity; specificity.

“İstatistik” verileri derleme, sınıflandırma, tablo ve grafiklerle özetleme, olasılık yardımıyla tasarımlama ve gözlem ilkelerini saptama, örneklem bilgilerinin anlamlılığını inceleme, yorumlama ve genelleme yöntemlerini veren bir bilimdir.<sup>[1]</sup> Kısa-cası istatistik, herhangi bir testin veya olayın nicel yönlerinin sayılarla anlatımıdır.

İstatistik yöntemleri diğer birçok bilimlerde olduğu gibi biyoloji, tıp, veterinerlik ve eczacılık bilimlerinde de yapılacak bilimsel araştırma ve değerlendirme çalışmalarında çok önemli yardımcı bir disiplindir. Tıpta ve eczacılıkta yeni bir tedavi yönteminin, yeni bir ilacın, yeni bir testin, ya da yeni bir ameliyat tekniğinin gerçek değeri ancak istatistik yöntemler yardımıyla en doğru biçimde belirlenebilir. İşte bu nedenlerle, tanınmış bir bilim

adamının dediği gibi: “Günümüzde biyolojik ve tıbbi problemleri tarafsız ve eleştireci bir düşünce içinde sayısal esaslara göre değerlendirmeye alışmamış bir hekim, bir bilimin temsilcisi değil, ancak bir sanatın uygulayıcısı olabilir.”<sup>[2]</sup>

Biyoistatistik, “canlıların özelliklerinin ve bu özellikleri etkileyen diğer tüm öğelerin saptanması, sayılması, düzenlenmesi ve değerlendirilmesi ile nedensel bağlantıları çözümlenmeye çalışan bilim dalı” olarak tanımlanabilir. Biyoistatistik iki ana biçimde kullanılmakta olup, bunlardan ilki verilerin toplanması ve dökümü, ikincisi ise çeşitli matematik yöntemlerle değerlendirilmesi şeklindedir.<sup>[3]</sup>

Dünya Sağlık Örgütü’nün tanımına göre tarama, uygulanması hızlı testlerle hastalık veya nor-

mal dışı özellikler taşıdığı o ana dek göze çarpmış olguların öncül belirlenmesidir.

Tarama ve tanı iki farklı kavramdır. Tarama ve tanıdaki ayrımlar kısaca “test” olarak adlandırılmaktadır. “Test”, yaptığı ayırım ve koyduğu tanı açısından güvenli olmalıdır. Başka bir deyişle testin iç ve dış geçerlilikleri yüksek olmalıdır.<sup>[4]</sup>

Tarama testleri, görünürde sağlıklı olan ancak belli bir hastalığı taşıması olası bireylerle, bu hastalığı taşımayan bireyleri ayırmaya olanak sağlamalıdır. Amaçları tanı koymak değildir. Taramada özellikle kesin sağlıklı olanların belirlenmesi önem taşımaktadır.

Tanıda, kuşku ve kesin hasta için ayırlama aşaması gerekmektedir. Bu aşamada tanıyı kesinleştirecek, tedaviye alınacak hastaları ayırmaya yönelik yöntemlerden geçeceklerdir.<sup>[4]</sup>

Tanı testleri görece olarak değerlendirilir. Testin özellikleri ölçüt yöntemlerinin sonuçlarına dayalı hesaplamalarla elde edilir. Özellikler hasta/sağlam ve test (+)/test (-) birlikte ele alınarak değerlendirilir. Tanı ve tarama testlerinin değerlendirilmesinde genellikle kullanılan ölçütler duyarlılık (*sensitivity*), özgüllük (*specificity*), genel doğruluk (*accuracy*), artı yorum gücü (*positive predictive value*) ve eksi yorum gücü (*negative predictive value*) kullanılır. Bu ölçütler için dört gözülü tablodan yararlanılarak testlerin sınır değerleri belirlenebilir (Tablo 1).

a) *Duyarlılık (sensitivity) (%)*: Gerçek hastaları ayırabilme yeteneği

$$a / (a + c) = GP / (GP + YN)$$

b) *Özgüllük (%) (specificity)*: Gerçek sağlamları ayırabilme yeteneği

$$d / (b + d) = GN / (GN + YP)$$

c) *Artı yorum gücü (%) (+ predictive value)*: Test pozitifler içerisinde doğruluk oranı

$$a / (a + b) = GP / (GP + YP)$$

d) *Eksi yorum gücü (%) (- predictive value)*: Test negatifler içerisinde doğruluk oranı

$$d / (c + d) = GN / (GN + YN)$$

e) *Genel doğruluk (%) (accuracy)*: Testin toplam doğru tanı koyma oranı

$$(a + d) / (a + b + c + d) = (GP + GN) / (GP + GN + YP + YN)$$

Klinikte rutin kullanılan tanı testlerinin evrensel alt-üst sınır (normallik sınırları) değerleri belirlenmiştir. Bu sınırlar kullanılan kimyasal reaktiflerin özelliklerine ve testin yapıldığı ısı derecesine göre çok az farklılık gösterir. Bazı sınırlar dünyada standardize edilmiştir (örn.: glikoz, karaciğer enzimleri, hemoglobin miktarı vb.). Buna karşılık tanıda kullanılmak üzere pek çok yeni test kullanıma sunulmaktadır. Fakat bu testlerin henüz standardize alt-üst sınırları belirlenmemiştir. Bu nedenle testlere ait sınırları saptamak için her laboratuvarın kendi sınırlarını belirlemesi gerekmektedir.

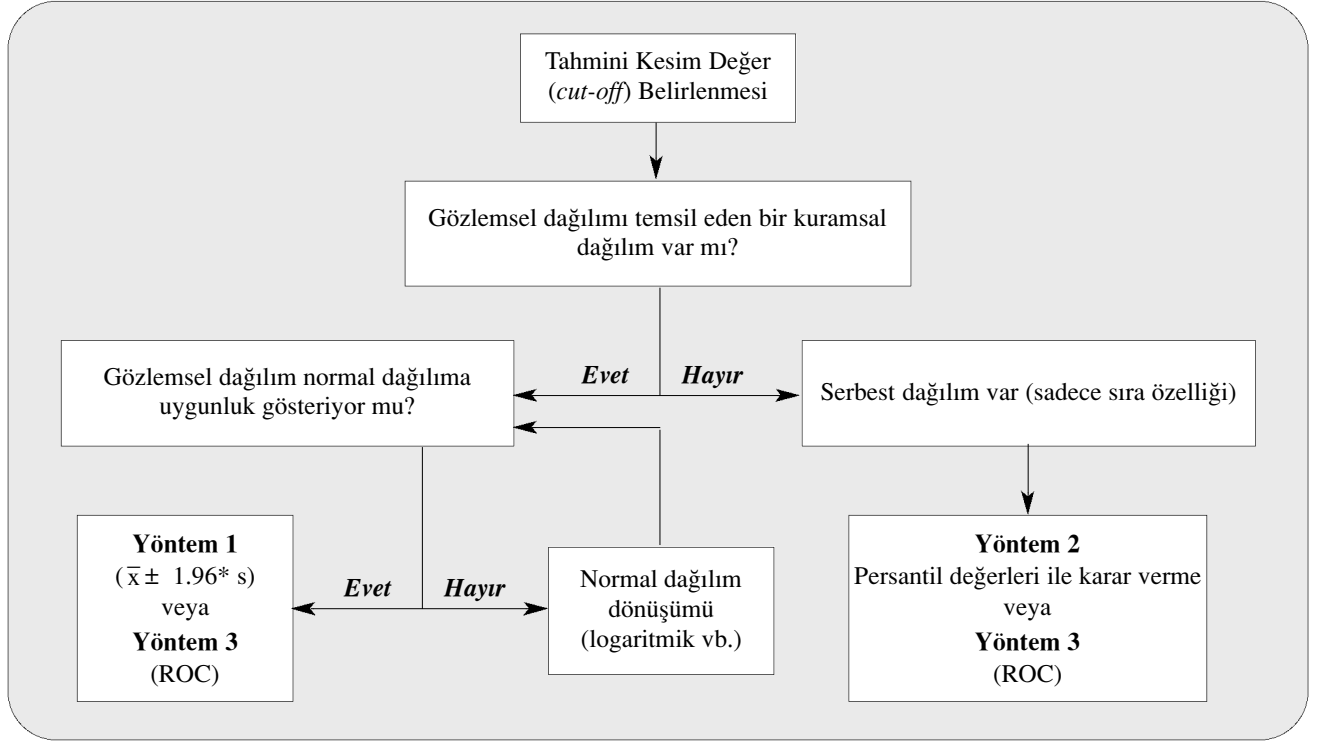
Bir testin alt-üst sınırlarının belirlenmesi için sağlıklı toplumdan rastgele seçilmiş deneklerin sayısı ana yığını temsil edecek büyüklükte olmalıdır.<sup>[5]</sup> Elde edilen test sonuçlarının tanımlayıcı istatistikleri (aritmetik ortalaması, standart sapması,

**Tablo 1**

Dört gözülü tablo

	Gerçek hasta	Gerçek sağlam	
Test sonucu	(+)	(-)	<i>Toplam</i>
Test (+)	a (GP)	b (YP)	a+b
Test (-)	c (YN)	d (GN)	c+d
<i>Toplam</i>	a+c	b+d	a+b+c+d

GP: Gerçek pozitif; YP: Yalancı pozitif; YN: Yalancı negatif; GN: Gerçek negatif.



Şekil 1. Kesim (tahmini referans) değerlerinin belirlenmesinde izlenecek yol.

ortanca değeri, persantileri ile en düşük ve en yüksek değeri) hesaplanır.<sup>[6]</sup> Hasta seçimi belirlendikten sonra dağılım sınanarak normal dağılıma uymayan gruplarda normal dağılımı elde edebilmek için istatistiksel yöntemlerle dönüşüm yapılmalıdır.<sup>[7]</sup> Laboratuvar test sonuçlarının ışığında aşağıdaki diyagrama göre alt-üst sınır değer belirleme yöntemlerinden biri seçilir (Şekil 1).

### Sınır Değerlerin Saptanma Yöntemleri

Sınır değerlerin saptanma yöntemleri, “*Macrophage Migration Inhibitory Factor*” (MIF) tümör marker değişkeninin (sürekli değişken) kolon

kanseri tanısında kullanımına ilişkin bir örnek verilerek açıklanacaktır.

Kolon kanseri tanısı konmuş 35 hasta ve 20 sağlıklı kişide (kolon kanseri olmayan) MIF değerlerinin ölçüldüğünü ve elde edilen değerlerin aşağıdaki tablolarda (Tablo 2, 3) özetlendiği varsayalım.

Bu örnek üzerinden üç ayrı yöntem irdelenecektir.

#### Yöntem 1. ( $\bar{x} \pm 1,96* sd$ )

MIF değişkeninin kontrol ve hasta gruplarında

Tablo 2

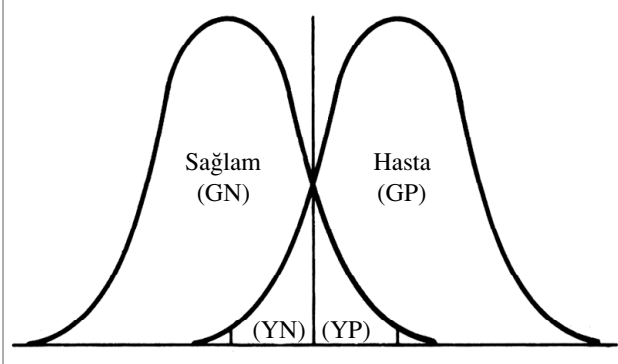
Hasta grubu (Kolon kanseri hastaları)

Hasta	(MIF)	ng/mL
55	11	15
57,5	11	15,5
8	11,5	17
9	11,5	17,5
10	12	17,5

Tablo 3

Kontrol grubu (Kolon kanseri olmayan)

Kontrol	(MIF)	ng/mL
15	16,4	9
8	10,8	10,4
13,7	8,2	9,6
9,6	21,2	15,1
13,5	11,5	12



Şekil 2. Hasta ve kontrol test değerlerinin kuramsal dağılımı.

normal dağılıma uygun olup olmadığı “Kolmogorov-Smirnov tek örneklem” testi ile sınanır. Sınama sonucunda kontrol ve hasta grubunun normal dağılıma uygunluk gösterdiği belirlenirse kontrol grubunda teste ait “aritmetik ortalama + (1,96) x standart sapma” ifadesi yardımıyla testin üst sınırı belirlenir. Hasta MIF verileri ile “aritmetik ortalama - (1,96) x standart sapma” ifadesi kullanılarak testin alt sınırı saptanır (Şekil 2).

Örneğimizde, MIF değişkeni “Kolmogorov-Smirnov tek örneklem” testi sonucunda kontrol grubunun normal dağılıma uygunluk göstermektedir ( $p=0,983$ ). Fakat hasta grubunun normal dağılıma ( $p=0,007$ ) uygunluk göstermediği görülmektedir. Bu nedenle, logaritmik dönüşüm yapılması gerekir. MIF değişkeni logaritması hesaplanarak “logMIF” değişkenine dönüştürüldü.

Logaritmik dönüşüm yapıldığında hem kontrol grubunun ( $p=1,0$ ), hem de hasta grubunun ( $p=0,222$ ) “logMIF” değişkeninin normal dağılımı gözlenir. Daha sonra kontrol grubunun “logMIF” değişkenine ait “aritmetik ortalama + (1,96) x standart sapma” ve hasta grubuna ait “aritmetik

ortalama - (1,96) x standart sapma” ifadeleriyle tahmini kesim (tahmini referans) değerlerinin bulunduğu aralık belirlenir. Bu sınırlar civarındaki test değerlerine ait duyarlılık ve özgüllükler hesaplanır. Duyarlılık ve özgüllüğü aynı anda yüksek olan test değeri, kesim (tahmini referans) değeri olarak belirlenir.

### MIF değerleri

Kontrol: Aritmetik ortalama ( $\bar{x}$ ): 12,43 ng/mL

Standart sapma (sd): 3,20

Hasta: Aritmetik ortalama ( $\bar{x}$ ): 25,21 ng/mL

Standart sapma (sd): 22,08

Üst sınır:  $\bar{x} + (1,96 \cdot SD) = 12,43 + (1,96 \cdot 3,20) = 18,7$  ng/mL (Kontrol grubundan)

Alt sınır:  $\bar{x} - (1,96 \cdot SD) = 25,21 - (1,96 \cdot 22,08) = 0$  ng/mL (Hasta grubundan)

Örnekte grupların birinde normal dağılım olmadığı için,

### LogMIF değerleri

Üst sınır:  $\bar{x} + (1,96 \cdot SD) = 1,0814 + (1,96 \cdot 0,1074) = 1,2919$  ng/mL; 10 tabanına göre antilog değişimi = 19,58 bulunur. (Kontrol grubundan)

Alt sınır:  $\bar{x} - (1,96 \cdot SD) = 1,2906 - (1,96 \cdot 0,2899) = 0,7224$  ng/mL; 10 tabanına göre antilog değişimi = 5,28 bulunur. (Hasta grubundan)

Bulunan bu kesim (tahmini referans) değerleri olan 5,28 ( $\approx 5$ ) ile 19,58 ( $\approx 20$ ) arasında 4 kesim (tahmini referans) noktası için duyarlılık ve özgüllükler hesaplanır. Kesim (tahmini referans) noktaları aşağıdaki gibidir (Tablo 4).

Tablo 4

Tahmini kesim noktalarının duyarlılık ve özgüllükleri

Kesim (tahmini referans) değeri (pozitiflik alt sınırı)	Duyarlılık (%)	Özgüllük (%)
$\geq 5$	100	0
$\geq 10$	85,71	25
$\geq 15$	62,86	80
$\geq 20$	34,29	95

**Tablo 5**

Tahmini kesim noktalarının duyarlılık ve özgüllükleri

Kesim (tahmini referans) değeri (pozitiflik alt sınırı)	Duyarlılık (%)	Özgüllük (%)
≥10	85,71	25
≥11	85,71	35
≥12	71,43	55
≥13	68,57	65
≥14	65,71	75
≥15	62,86	80

Tablodan görüldüğü gibi tanı değerlerinin her ikisini en yüksek yapan sınır değer 10 - 15 aralığında olduğu görülmektedir. Bu aralığı daha ayrıntılı incelersek Tablo 5'deki sonuçları elde ederiz.

En iyi kesim değerinin  $\approx 14$  ng/mL olduğu görülmektedir.

### Yöntem 2.

Sınır değerlerin belirlenmesinde testin hasta ve kontrol sonuçlarının ortanca (50. persantil) değeri baz alınarak 2,5. - 97,5. persantil değerleri, persantil tablosu kullanılarak hesaplanır.<sup>[9]</sup> 2,5 ve 97,5. persantil değerleri temsil edebilecek gözlem sayısı yoksa 5. - 95. veya 25. - 75. persantil değerleri gibi alt-üst sınırlar kullanılabilir. Kontrol grubunun maksimum değerine yakın değerler içinden belirlenen kesim değeri üst sınır; hasta grubunun minimum değerine yakın değerler arasından alt sınır belirlenir.<sup>[10]</sup> Örnekte her ne kadar dönüşüm sonrası normallik koşulu gerçekleşmiş olsa da, yöntem iki aynı örnek üzerinden açıklanacaktır.

Örnekteki kontrol grubunun MIF değerlerine göre:

90. persantil: 16,33 ng/mL

95. persantil: 20,96 ng/mL

97,5. persantil: 21,20 ng/mL

Örnekteki hasta grubunun MIF değerlerine göre:

2,5. persantil: 8,0 ng / mL

5. persantil: 8,8 ng/mL

10. persantil: 10,0 ng/mL

Bulunan bu kesim değerleri olan 8,0 ile 21,20 ( $\approx 22$ ) arasında 4 kesim (tahmini referans) noktası için duyarlılık ve özgüllükler hesaplanır. Tahmini kesim noktaları Tablo 6'daki gibidir. Tabloda görüldüğü gibi en uygun kesim değerinin  $\approx 13$  ng/mL olduğu görülmektedir.

### Yöntem 3.

ROC (*Receiver Operating Characteristics*) eğrisinin çözümlene yönteminde, belirli bir tanım aralığında sürekli değerler alan bir değişkenin (sürekli değişken) tanı testi olarak kullanımı amaçlanmaktadır.<sup>[11]</sup> *Receiver Operating Characteristic*

**Tablo 6**

Tahmini kesim noktalarının duyarlılık ve özgüllükleri

Kesim (tahmini referans) değeri	Duyarlılık (%)	Özgüllük (%)
8	100	5
13	68,57	65
18	37,14	95
22	31,43	100

**Tablo 7**

Hasta grubu sıralı değerler

Hasta	(MIF)	ng/mL				
8	11	12	15,5	17,5	23	55
9	11	13	17	18	27,5	57
10	11	14	17	19	36,3	57,5
10	11,5	15	17,5	20	38	93
10	11,5	15	17,5	22,5	46,5	95

(ROC) eğrisinden yararlanılarak sınır değerlerinin belirlenmesinde istatistik paket program veya grafik çizerek hesaplama yöntemi kullanılır. Bu eğrinin çizimi için kesim (tahmini referans) noktalarına ihtiyaç vardır.

ROC çözümleme yöntemini, yukarıdaki örneklerle açıklayalım.

Örneğimizdeki hasta ve kontrol grubuna ait MIF değerlerini küçükten büyüğe sıraya dizilerek sıralı değerlerden oluşan tablolar elde edilir (Tablo 7, 8).

Hasta ve kontrol gruplarına ait tanımlayıcı istatistikler aşağıdaki gibi elde edilir.

Çeşitli pozitiflik sınır değeri için (*cut-off values*, kesim değerleri) duyarlılık ve  $(1 - \text{özgüllük})$  değerleri hesaplanır.

Hasta ve kontrol grupları birlikte ele alındığında, en küçük değer 8,0 en büyük değer ise 95,0 olarak belirlenir, 8,0 ile 95,0 arasında en az 10 pozitiflik sınır değeri tesbit edilir. Örnek için 8, 10, 12, 14, 16, 18, 20, 25, 30 ve 40 pozitiflik sınır de-

**Tablo 8**

Kontrol grubu sıralı değerler

Kontrol	(MIF)	ng/mL		
8	10,4	12		15
8,2	10,8	12,7		15,1
9	11,5	12,8		15,7
9,6	11,6	13,5		16,4
9,6	11,7	13,7		21,2

ğerleri kullanılmaktadır. Kesim (tahmini referans) değerlerinin sayısının artması ROC eğrisinin duyarlılığını arttırmaktadır.

Pozitiflik sınır değeri (*cut-off value*) 16 için duyarlılık ve özgüllük değerlerini hesaplamak için:  $MIF \geq 16$  ise test sonucu pozitif (kişi hasta) aksine  $MIF \leq 16$  ise test sonucu negatif (kişi sağlıklı) kabul edilecektir. Hasta grubunda 19 kişide, kontrol grubunda ise 2 kişide  $MIF \geq 16$ 'dır.

16 sınır değeri için; doğru pozitif, yanlış pozitif, yanlış negatif ve doğru negatif sayıları tabloda özetlenmiştir (Tablo 9).

$$\text{Duyarlılık} = 19 / 35 = 0,54$$

$$\text{Özgüllük} = 18 / 20 = 0,90$$

$$(1 - \text{özgüllük}) = 1 - 0,90 = 0,10$$

16 sınır değeri için (0,10; 0,54) noktası elde edilir.

ROC eğrisinin çizileceği noktalar, diğer tüm sınır değerleri için benzer işlemleri tekrarlayarak belirlenir.

**Tablo 9**

Örneğe ait dört gözlü tablo

Test sonucu	Hasta	Sağlam	Toplam
	Kolon kanseri	Kolon kanseri	
MIF (+)	19 DP	2 YP	21
MIF (-)	16 YN	18 DN	34
Toplam	35	20	55

**Tablo 10**

Sınır değerleri için elde edilen duyarlılık, özgüllük ve (1 – özgüllük) değerleri

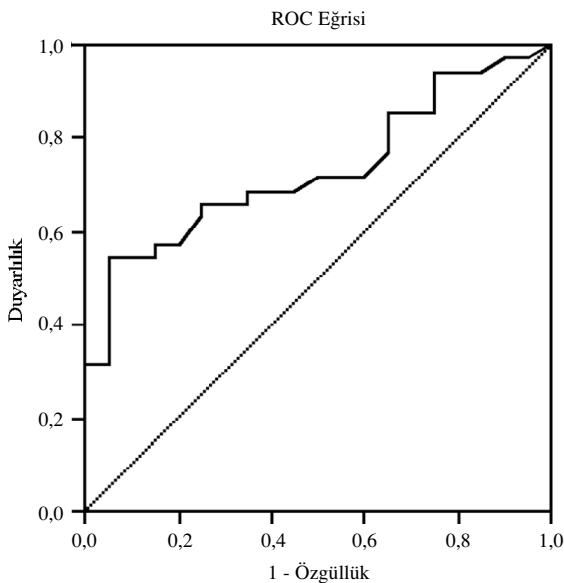
MIF pozitiflik sınır değeri ( $\geq$ )	Duyarlılık DP	%	Özgüllük DN	%	(1- özgüllük) %
8	35	100	0	0	100
10	33	94	5	25	75
12	25	71	10	50	50
14	23	66	15	75	25
16	19	54	18	90	10
18	14	40	19	95	5
20	12	34	19	95	5
25	9	26	20	100	0
30	8	23	20	100	0
40	6	17	20	100	0

Duyarlılık (y) ekseninde, (1 – özgüllük) değerleri ise (x) ekseninde temsil edilir.

Tüm sınır değerleri için elde edilen duyarlılık, özgüllük ve (1 – özgüllük) değerleri Tablo 10'da verilmektedir.

10 ayrı pozitiflik sınır değeri için elde edilen duyarlılık ve (1 – özgüllük) koordinat değerlerini kullanarak ROC eğrisi çizilir (Şekil 3).

MIF tümör marker değişkeninin tanı değerinin



Şekil 3. MIF tanı testine ait ROC eğrisi.

gücü (hasta ve sağlıklı ayırt edebilme gücü) ROC eğrisinin altında kalan alan ile ifade edilir. Alan değeri 1'e yaklaştıkça tanı değeri yükselir. %100'lük tanı gücünde alan değeri 1'e eşit olur.

Yukarıdaki tablodan görüldüğü gibi en uygun kesim değeri  $\approx 14$  ng/mL olduğu görülmektedir.

### Kaynaklar

1. İstatistik terimleri sözlüğü. Türk Dil Kurumu Yayınları 510, s. 59.
2. Low CR. Symposium sur l'Enseignement de la statistique aux Etudiants en Medicine en Europe, 21-23 Aout, Geneve: Publications de l'O.M.S; 1962.
3. Şenocak M. Biyoistatistik. 2. baskı, İstanbul: İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi; 1998.
4. Sackett DL. The usefulness of laboratory tests in health-screening programs. Clin Chem 1973;19(4):366-72.
5. Elveback LR, Guillier CL, Keating FR Jr. Health, normality, and the ghost of Gauss. JAMA 1970;211(1):69-75.
6. Solberg HE. The theory of reference values Part 5. Statistical treatment of collected reference values. Determination of reference limits. J Clin Chem Clin Biochem 1983;21(11):749-60.
7. Solberg HE. Statistical treatment of reference values. Bull Molec Biol Med 1983;8:13-9.
8. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839-43.

9. Letner C. Introduction to statistics. Statistical tables, Geigy Scientific tables, Ciba-Geigy Ltd. Basel-Switzerland, 1992.
10. Shapiro DE. The interpretation of diagnostic tests. Stat Methods Med Res 1999;8(2):113-34.
11. Obuchowski NA, Lieber ML, Wians FH Jr. ROC curves in clinical chemistry: uses, misuses, and possible solutions. Clin Chem 2004;50(7):1118-25.