

Tanı testlerinde performans değerlendirilmesi: temel kavramlar

Evaluation of the performance of the diagnostic tests: Basic concepts

Muzaffer BAŞAK, Mehmet ERTÜRK, Alper ÖZEL

Şişli Etfal Eğitim ve Araştırma Hastanesi 1. Radyoloji Kliniği

GİRİŞ

Günümüz tıbbında teşhis süreçlerinde tanı testlerinin önemi tartışılmaz. Bu testler ile çok temel olarak bir hastalık ile karşı karşıya olup olmadığımızı sorgulamaya çalışırız. Testin pozitif olması o kişide söz konusu hastalığın bulunduğunu, negatif olması ise hastalığın bulunmadığını gösterir. Tanı testlerinin performansları ise bazı kavramlar ile belirlenmeye çalışılır. Bu kavramların en sık kullanılanları ise duyarlılık (sensitivity), özgünlük (specificity), pozitif kestirim değeri (positive predictive value), negatif kestirim değeri (negative predictive value) ve kesinlik (accuracy) olarak adlandırılanlardır (1).

TANIMLAR VE ÖRNEKLER

Bu noktada bazı tanımlamalar yapmamız gerekmektedir. Gerçekten hasta olan ve testin de pozitif sonuç verdiği olgular gerçek pozitif (true positive=tp) olarak adlandırılır. Gerçekten sağlıklı olan ve testin de negatif sonuç verdiği olgular ise gerçek negatif (true negative=tn) olgular olarak sınıflandırılırlar. Hasta olmayan ama testin hatalı olarak pozitif olarak belirlediği olgular yalancı pozitif (false positive=fp), hasta olmasına rağmen test tarafından yine hatalı olarak negatif olarak değerlendirilen olgular ise yalancı negatif (false negative=fn) sonuçları oluştururlar (2,3,4).

Yazışma Adresi:

Muzaffer Başak,
Konaklar Mah. Şebboy Sok. Petekler Sitesi H Blok D:17
80620 Yenilevent – İstanbul / Türkiye
Tel: 0212-324 05 44 Fax: 0212-512 63 62

Örneğin akut apendisit şüphesi ile hastaneye başvuran hastaları ele alalım. Bu hastaların ultrason ile incelenmelerinde sağ alt kadranda komprese edilemeyen ve 6 milimetereden fazla bir çapa sahip tübüler yapı izlenmesinin akut apendisit anlamına geleceğini iddia edebilir ve bu iddiamızı da postoperatif sonuçlar ile doğrulamaya çalışabiliriz. Diyelim ki, 100 hastayı ultrason ile muayene ettik ve yukardaki kriterlerden hareketle 60'ına akut apendisit teşhisi koyduk, 40'ında ise akut apendisit olmadığını raporladık. Ve yine diyelim ki cerrahi departmanı kendi klinik kanılarına göre hareket etti ve bu 100 hastayı da ameliyata aldı. Ameliyat sonrasındaki patolojik değerlendirmede, bizim akut apendisit olarak değerlendirdiğimiz 60 hastanın 50'sinin gerçekten akut apendisit olduğu, kalan 10 tanesinde ise böyle bir durumun söz konusu olmadığı, sağlıklı olarak değerlendirdiğimiz 40 hastanın ise 8 tanesinin patolojik olarak akut apendisit tanısı aldığını kabul edelim.

Bu durumda 50 gerçek pozitif, 10 yalancı pozitif, 32 gerçek negatif ve 8 yalancı negatif sonuç ortaya çıkmış olacaktır. Şimdi bazı sorular sorabiliriz:

1. Hasta bir kişinin pozitif bir test sonucu almasının olasılığı ne kadardır?

Bu olasılığa duyarlılık adı verilir ve şöyle hesaplanır (2,3,4):

$$\begin{array}{lcl} \text{Gerçek pozitif sonuçlar} & & tp \\ \hline \text{Tüm hastalar} & = & tp + fn \end{array}$$

Burada önemli olan nokta tüm hastaların sayısının gerçek pozitif olgular ile yalancı negatif olguların sayısının toplamına eşit olduğunun kavranmasıdır. Gerçek pozitifler hasta olup da testin de pozitif sonuç verdiği, yalancı negatifler ise hasta olmalarına rağmen testin negatif sonuç verdiği olgulardır ve bu iki grubun toplanması bize incelediğimiz popülasyondaki tüm hastaların sayısını vermektedir.

O halde örneğimizde duyarlılık = $\frac{50}{50+8} = 0.86$ olarak hesaplanacaktır.

2. Sağlıklı bir kişinin negatif test sonucu alması olasılığı ne kadardır?
Bu olasılığa özgünlük adı verilir ve şöyle hesaplanır (2,3,4):

Gerçek negatif sonuçlar tn
----- = -----
Tüm sağlıklılar $tn + fp$

Burada önemli olan nokta tüm sağlıklıların sayısının gerçek negatif olgular ile yalancı pozitif olguların sayısının toplamına eşit olduğunun kavranmasıdır. Gerçek negatifler sağlıklı olup da testin de negatif sonuç verdiği, yalancı pozitifler ise sağlıklı olmalarına rağmen testin pozitif sonuç verdiği olgulardır ve bu iki grubun toplanması bize incelediğimiz popülasyondaki tüm sağlıklıların sayısını vermektedir.

O halde örneğimizde özgünlük = $\frac{32}{32+10} = 0.76$

3. Pozitif test sonucu olan bir kişinin gerçekten hasta olma olasılığı ne kadardır?
Bu olasılığa pozitif kestirim değeri adı verilir ve şöyle hesaplanır (4,5):

Gerçek pozitif sonuçlar tp
----- = -----
Tüm pozitif sonuçlar $tp + fp$

O halde örneğimizde pozitif kestirim değeri = $\frac{50}{50+10} = 0.83$ olarak hesaplanacaktır.

4. Negatif test sonucu olan bir kişinin gerçekten sağlıklı olma olasılığı ne kadardır?
Bu olasılığa negatif kestirim değeri adı verilir ve şöyle hesaplanır (4,5):

Gerçek negatif sonuçlar tn
----- = -----
Tüm negatif sonuçlar $tn + fn$

O halde örneğimizde negatif kestirim değeri = $\frac{32}{32+8} = 0.80$ olarak hesaplanır.

5. Kesinlik ise doğru test sonuçlarının test edilen tüm insanların sayısına bölünmesi ile hesaplanan bir büyüklüktür:

Gerçek pozitif sonuçlar + gerçek negatif sonuçlar $tp + tn$
----- = -----
Test edilen tüm insanlar $tp + tn + fp + fn$

O halde örneğimizde kesinlik = $\frac{82}{100} = 0.82$

Şimdi bu kavramları biraz daha açmaya çalışalım. Duyarlılığın gerçekten hasta bir kişinin pozitif test sonucu alma olasılığı olduğunu söylemiştik. Örneğimizde bu olasılık %86 olarak belirlendi. Başka bir deyişle 100 gerçekten hasta kişinin 86'sı test sonucunda "hasta" olarak çıkmaktadır. 14 kişi ise hasta olmasına rağmen test sonucunda "sağlam" olarak sınıflandırılacaktır. Eğer bu değer örneğin %20 olsa idi, gerçekten hasta olanların sadece %20'sinde pozitif test sonuçları elde edilecekti. Özgünlük ise gerçekten sağlıklı bir kişinin negatif test sonucu alma olasılığı idi. Örneğimizde bu

olasılık %76 olarak belirlenmişti. Yani 100 sağlam kişiye ultrason ile bakıldığında 76'sı negatif çıkarken, 24'ü yanlış yere akut apendisit tanısı alacak demektir. Bu iki değer genelde birbirleri ile ters bir ilişki içindedir. Duyarlılık arttığında özgünlük azalırken, özgünlüğün artması ise duyarlılığın azalması sonucunu doğurabilir. Pozitif kestirim değerinin ise pozitif test sonucu olan bir olgunun gerçekten hasta olma olasılığıdır. Örneğimizde bu değer %83'tü. Yani pozitif bir test sonucu %83 olasılıkla hastalık anlamına gelmektedir. Negatif kestirim değeri ise negatif test sonucu olan bir olgunun gerçekten sağlıklı olma olasılığıdır. Örneğimizde bu değer %80'di. Bunun anlamı negatif bir test sonucunun %80

kesinlik, pozitif kestirim değeri ve negatif kestirim değerini hesaplayalım. Öncelikle test ve kültür sonuçlarını birlikte gösterdiğimiz bir tablo oluşturmamız gerekecektir (tablo 1).

Örneğimizde altın standart kültür olarak seçildiğinden, test sonuçları kültür sonuçları ile karşılaştırılarak gerçek ve yalancı sonuçlar belirlenecektir. Kültür sonucunun ve test sonucunun pozitif olduğu hastalar gerçek pozitif olgularımızdır. Bu hastalarda test yanılmamıştır. Kültür sonucu negatif olmasına rağmen test sonucu pozitif olan hastalar ise yalancı pozitiflerdir. Aynı şekilde kültür sonucu negatif olan hastalardan test sonucu da negatif olanlar gerçek negatif olgularımız, test sonucu pozitif olanlar ise yalancı pozitiflerimiz olarak sınıflandırılmıştır.

Tablo 1: Örnek 2'nin sonuçları

Test	Kültür			
		Pozitif	Negatif	Toplam
Pozitif		90(tp)	40(fp)	130
Negatif		25(fn)	30(tn)	55
Toplam		115	70	185

olasılıkla sağlıklı birinden kaynaklanıyor olduğudur. Kesinlik ise duyarlılık ve özgünlük kavramlarının tek bir kavramda özetlenmiş hali diye düşünülebilir. Örneğimizdeki %82'lik kesinlik

Gerekli hesaplamaları yaptığımızda şu sonuçlar ortaya çıkar (tablo 2)

Bu değerler şöyle yorumlanabilir:

1. Duyarlılık %78'dir. Yani bu test 100 hastaya uygulandığında bunların 78 tanesinde pozitif sonuç

Tablo 2: Örnek 2'nin değerlendirilmesi

Duyarlılık	$=tp / (tp+fn)$	$=90/115 =0.78$
Özgünlük	$=tn / (tn+fp)$	$=40/70 =0.57$
Pozitif kestirim değeri	$=tp / (tp+fp)$	$=90/130 =0.69$
Negatif kestirim değeri	$=tn / (tn+fn)$	$=30/55 =0.54$
Kesinlik	$=(tp+tn) / (tp+tn+fp+fn)$	$=120/185=0.64$

değeri test edilen insanların %82'sinin hasta-sağlıklı bağlamında doğru sınıflandırıldığını göstermektedir.

Anlatılanları bir başka örnek üzerinde uygulayalım:

Tüberküloz ile ilgili yeni bir serolojik test deneniyor olsun. Test sonuçları kültür sonuçları ile karşılaştırılarak testin duyarlılık, özgünlük,

verecek, 22 hasta ise hatalı olarak sağlam olarak değerlendirilecektir.

2. Özgünlük %57'dir. Yani bu test 100 sağlıklı insana uygulandığında bunların 57 tanesinde negatif sonuç verecek, 43 kişi ise hasta olarak değerlendirilecektir.

3. Pozitif kestirim değeri %69'dur. Yani bu testin pozitif olması %69 olasılık ile hastalık

göstermektedir.

4. Negatif kestirim deaeri %54'tür. Yani bu testin negatif çıkması %54 olasılık ile hastalık yok anlamına gelmektedir.

5. Kesinlik %64'tür. Yani test uygulanan kişilerin %64'ü hasta ve sağlıklı olmak üzere doğru sınıflandırılmaktadır.

SONUÇ

Bu yazıda tanı testlerinin performansını tarif etmekte çok kullanılan bazı kavramları açıklamaya çalıştık. Bu kavramların bilinmesinin testlerin doğru seçilmesi ve değerlendirilmesinde ve sonuç olarak da ayırıcı tanı ve tedaviye yönelik karar süreçlerinde çok önemli olduklarına inanıyoruz.

KAYNAKLAR

1. Doubilet MP: Statistical techniques for medical decision making: applications to diagnostic radiology. AJR, 150:745-750, 1988.
2. Hayran M, Özdemir O: Tıpta Karar Verme. Bilgisayar, İstatistik ve Tıp, İkinci baskı, Hekimler Yayın Birliği, Ankara, 1996, 453-460.
3. Knapp RG, Miller MC: Describing the performance of a diagnostic test. Clinical epidemiology and biostatistics, Harwall Publishing Company, Malvern Pennsylvania, 1992, 34-39.
4. Armitage P, Berry G, Matthews JNS: Statistical methods in epidemiology. Statistical methods in medical research, Blackwell science, 693-698.
5. Köbberling J: Der praediktive Wert diagnostischer Massnahmen. Dtsch. Med. Wschr., 107:591-595, 1982.