



LABORATUVAR TESTLERİNİN KLİNİK PERFORMANSI

Türk Biyokimya Derneği

Doğan Yücel

Ankara
2021

Açıklama

Tıbbi laboratuvar testlerinin klinik performansının deęerlendirilmesine yönelik bu kısa yazı, daha önce verilen eğitimlerden derlenerek 2015 yılında hazırlanmış, 2021 yılında küçük deęişikliklerle Türk Biyokimya Derneęi sitesine eklenmiştir.

Doęan Yücel

LABORATUVAR TESTLERİNİN KLİNİK PERFORMANSI

Tıp ve laboratuvar ortamındaki dinamizm hızla sürmekte ve tıp ortamında laboratuvarın önemi her geçen gün artmaktadır. Klinisyenler her gün hastalarından aynı anda onlarca test istemektedir. Günümüzde klinik karar oluşumunda tıbbi laboratuvarın rolünün, henüz kanıtlanamasa da, %70-80 oranına ulaştığı öne sürülmektedir. Dinamik laboratuvar teknolojisi, tıbbi laboratuvar ortamına her geçen gün yeni testlerin girmesine kapı açmakta, test menüsü gittikçe büyümektedir. Aynı zamanda kullanılan testlerin karmaşıklığı da artmaktadır. Klinisyenlerin bu test yoğunluğunu ve karmaşıklığını takip edebilmesi mümkün değildir. Özellikle test seçiminde ve test sonuçlarının yorumunda laboratuvar disiplinlerinin desteği kaçınılmaz hale gelmiştir. Bu durum laboratuvar uzmanının sorumluluk alanını da genişletmektedir. Artık laboratuvarcı sadece laboratuvarında gelen istemlerin analizini sağlayan kişi değildir. Aynı zamanda gerek klinisyenlerce laboratuvar istemlerinin doğru yapılması ve örneklerin laboratuvara taşınması, laboratuvar hatalarının azaltılması ve hasta güvenliği, gerekse analiz sonrasında test sonuçlarının doğru yorumlanması ve doğru kullanılması, artık laboratuvar uzmanının görev alanına girmektedir. Ek olarak kullanılan testin uygulanabilirliği, sonuçlandırılma zamanı ve maliyeti de bilinmesi gereken özelliklerdir.

Konuyla ilgili bazı kavramlar aşağıda verilmektedir.

Tanısal Test (Diyagnostik Test). İnsan organizması kaynaklı materyalin, insanda sağlık durumunun değerlendirilmesi ya da herhangi bir hastalığın veya bozukluğun tanısı, önlenmesi veya tedavisi hakkında bilgi sağlanması amacıyla incelenmesidir.

Laboratuvar Testlerinin Kullanımı. Laboratuvar testleri pek çok amaç için kullanılır. Başlıca klinik kullanım alanları tarama/risk belirleme, tanı ve takip olarak özetlenebilir.

Tanı, semptomları olan hastaya laboratuvar testleri uygulayarak söz konusu hastalığın tanımlanmasını amaçlar.

Tarama, sağlıklı, asemptomatik kişide herhangi bir hastalığın veya defektin saptanması ve kişide yol açacağı riskin belirlenmesidir.

Takip ise laboratuvar testlerinin zaman içinde tekrarlanarak hastanın durumunun değerlendirilmesidir. Zaman dilimi kısa (örneğin akut hastalık sırasında hastaneye yatırılan hastanın takibi), orta uzunlukta (örneğin reküransı takip etmek için tümör belirteçlerinin izlenmesi) veya uzun (diyabetli hastada HbA1c ile gliseminin takibi) tutulabilir.

Laboratuvardan test istenme nedenleri genel olarak şöyledir: Asemptomatik kişide hastalık tarama: %30; Tanı: %37; Tedavi takibi: %33.

Klinik laboratuvar testleri, bunların dışında veri analizi, eğitim ve araştırma gibi amaçlarla da kullanılır

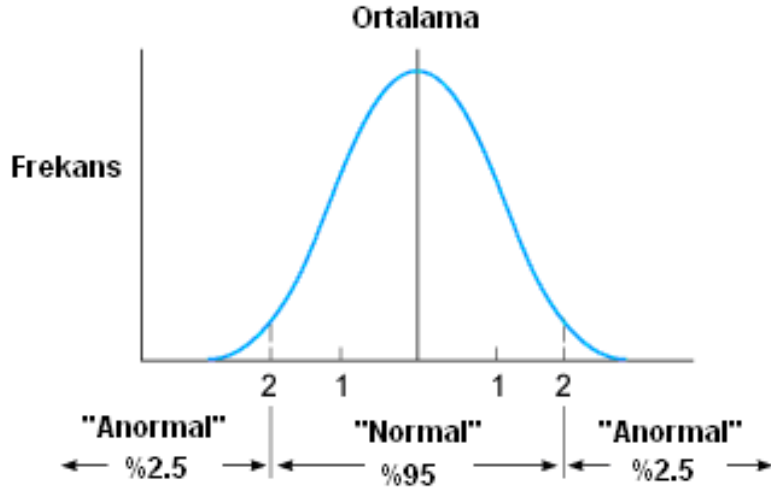
Analitik Performans ve Klinik Performans. Tanısal testlerin analitik ve klinik performansları farklı kavramlardır. Ölçütleri de farklıdır. Analitik performansın ölçütleri **doğruluk** (accuracy, şimdi trueness), **geri kazanım** (recovery), **kesinlik** (çalışma içi, çalışmalar arası, gün içi, günler arası, cihaz içi, laboratuvar içi, ara kesinlik ve total kesinlik), **tekrarlanabilirlik** (repeatability), **yeniden üretilebilirlik** (reproducibility), **rapor aralığı** (reportable range), **analitik duyarlılık** ve ilgili kavramlar (limit of blank, limit of detection, limit of quantitation), **analitik özgüllük** (interferanslar) gibi yöntem geçerliliğinin ölçütleri olan teknik özelliklerdir.

Herhangi bir tanısal testin analitik performansı çok iyi olabilir ancak klinik performans aynı ölçüde iyi olmayabilir. Analitik performans ile klinik performansın kesişme noktası bir bakıma referans aralıklarıdır.

Referans Aralık Kavramı. Test sayısı ve karmaşıklığının artışı klinisyenlerin, hem uygun test seçimini, hem de test sonuçlarını yorumlamasını güçleştirir. Bugün klinisyenler test yorumunda genel olarak popülasyona dayalı referans aralıkları kullanılmaktadır. Referans aralıklar sağlıklı popülasyonun %95'ini kapsar. İstatistiksel olarak (şansa bağlı olarak) herhangi bir test için sağlıklı bireylerin %5'inde referans sınırların dışında test sonucu elde etmek olasıdır. Aynı olasılık düzeyi göz önüne alındığında, sağlıklı kişiden iki test istendiğinde, her iki test sonucunun da referans sınırlar içinde kalma olasılığı $0.95 \times 0.95 = 0.9025$ veya %90.3'tür. Test sayısı 5'e çıktığında bu oran %77.4'e, 10'a çıktığında %59.9'a, 20 test istendiğinde %35.8'e düşer. Başka deyişle, sağlıklı bir kişiden 20 test istenirse bu test sonuçlarından birinin referans aralık dışında çıkma olasılığı %64'tür. Bu durum klinisyeni tanı için yeni arayışlara sokmakla kalmaz, hastayı da endişeye sevkeder. Dolayısıyla referans aralık kavramının klinisyenlerce iyi bilinmesi gereklidir.

Tanısal (klinik) performans ile ilişkili önemli bir kavram da **klinik karar sınırı** kavramıdır.

Klinik Karar Sınırı. Referans sınırlardan farklı bir kavramdır. Çünkü belirli spesifik bir tıbbi durum ile ilişkili ya da bu tıbbi duruma dayanan sınırlardır, klinik karar sınırları. Klinik karar sınırı, acil girişim gerektiren ya da tedavide değişiklik yapılmasına işaret eden "**Kritik Değerler**"i de kapsar. Veya herhangi bir klinik durum ya da hastalık için belirlenen kesim (cut-off) değerleri (eşik değerler) olabilir. İlaçların terapötik aralığı veya toksisite sınırları da bu kapsama girer.

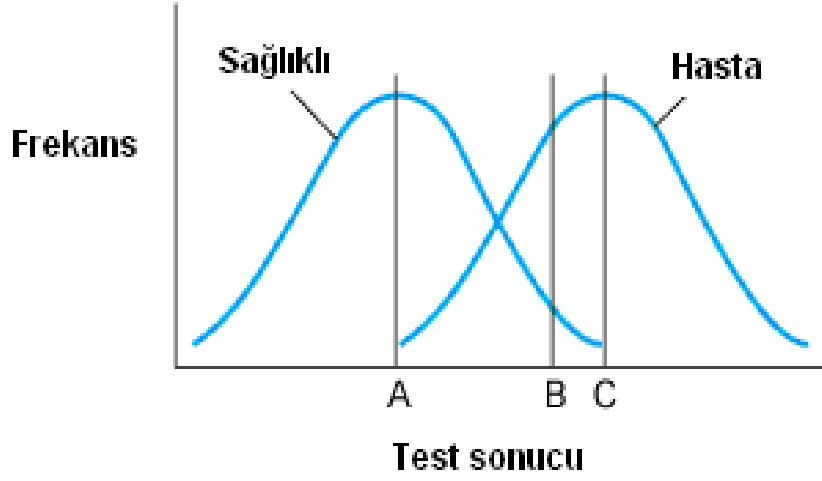


Şekil 1. Referans aralık kavramı. Referans aralık sağlıklı popülasyonun ortadaki %95’lik kesimini kapsar.

Tanısal Performans. Klinisyenlerce bilinmesi gerekli çok önemli diğer bir nokta testin tanısal (klinik) performansı veya tanısal doğruluğudur. Testin tanısal değeri ile ilgili temel veriler **Duyarlılık** (Sensitivite), **Özgüllük** (Spesifite), **Pozitif ve Negatif Öngörü Değerleri** (Prediktif Değer), **Pozitif ve Negatif Olabilirlik Oranları** (Likelihood Ratio), **Odds oranları** gibi klinik karar oluşumunda gerekli kavramlardır.

Duyarlılık, testin hastalığı belirleme gücünün göstergesidir; yüzde olarak ifade edilir. Test sonucunun hastalığı olanlarda pozitiflik oranıdır. Bir testin duyarlılığı örneğin %90 ise, bu durum, testin hastaların %90’ında pozitif sonuç verdiği (doğru pozitiflik: DP) anlamına gelir. Veya hastaların %10’unda yanlış negatiflik (YN) verdiğini belirtir. Duyarlılığı yüksek testler, genel olarak, tanıyı dışlamak için değerlidir. Çünkü yanlış negatif sonuç oranı düşüktür. Örneğin, HIV enfeksiyonunun dışlanması için Anti-HIV antikoru gibi duyarlılığı çok yüksek testler seçilmelidir.

Özgüllük, hastalığın olmadığını gösteren bir özelliktir ve yüzde olarak ifade edilir. Hasta olmayan kişilerde testin negatiflik yüzdesidir. Örneğin, testin özgüllüğü %90 ise, bu durum, testin hastalığı olmayan kişilerde %90 oranında negatif (doğru negatiflik: DN), %10 oranında yanlış pozitif (YP) olduğu anlamına gelir. Özgüllüğü yüksek bir test, tanıyı doğrulamak amacı için değer taşır, çünkü YP oranı çok düşüktür. Örneğin, HIV virüsü taramasında anti-HIV pozitif bulunan kişilerde Westernblot yöntemiyle doğrulama yapılır.



Şekil 2. Eşik değere göre duyarlılık ve özgüllüğün değişimi. Eşik olarak A değeri seçildiğinde duyarlılık yüksek (%100) ama özgüllük düşüktür; C’de özgüllük yüksek (%100) ama duyarlılık düşüktür; B’de ise referans aralık olarak kullanılan ortalama $\pm 2s$ sınırları esas alınmıştır.

Herhangi bir testin duyarlılık ve özgüllüğü, hasta ve normallerin ayrımında kullanılan analit değerine (eşik değeri) bağlıdır. Eğer testin eşik değeri düşük tutulursa duyarlılığı artırılır ama özgüllüğü düşer. Tersine eşik değeri yükseltirirse duyarlılık azalır ama özgüllük artar (Şekil 2).

Tanısal testlerin duyarlılık ve özgüllükleri ve bu değerlere bağlı olarak DP, YP, DN ve YN oranları en iyi **2x2 tablosu** ile görülür. 2x2 tablosunda tercih edilen yapı, hastalık varlığı veya yokluğunun düşey sütunlara konulması, yatay olarak pozitif ve negatif test sonuçlarının verilmesidir.

2x2 Tablosu

	Test Sonucu Pozitif	Test Sonucu Negatif
Hasta olanlar	DP	YN
Hasta olmayanlar	YP	DN

Veya

	Hastalık Var	Hastalık Yok
Pozitif Test	DP	YP
Negatif Test	YN	DN

- DP (Doğru Pozitif):** Testin hastalığı olanlar için pozitif sonuç vermesi (Gerçek Pozitif)
DN (Doğru Negatif): Testin hastalığı olmayanlarda negatif sonuç vermesi (Gerçek Negatif)
YP (Yanlış Pozitif): Testin hastalığı olmayanlarda pozitif sonuç vermesi
YN (Yanlış Negatif): Testin hastalığı olanlarda negatif sonuç vermesi

Genellikle bu ikinci tablo şekli tercih edilir ve kullanılır.

Tanısal duyarlılık ve özgüllüğün saptandığı en basit çalışmalar, çapraz kesitli çalışmalardır. Bu çalışmalarda hastalarla sağlıklı kontroller aynı teste tabi tutulur ve sonuçlar derlenir. Ancak kesitli çalışmalarının spektrum sapması nedeniyle tanısal doğruluğu yüksek gösterdiği bilinmektedir. Tanısal doğruluğun en iyi saptandığı çalışmalar **prospektif, randomize** klinik çalışmalardır. Bu çalışmalarda aynı semptomlara sahip hastalar, ardı ardına gerçek tanı bilinmeksizin teste tabi tutulur, daha sonra **“altın standardı”** tanı yöntemiyle hastanın kesin tanısı konulur ve sonuçlar buna göre derlenir. Tanısal doğruluk çalışmaları için önerilen çalışma yaklaşımları için Standards for Reporting of Diagnostic Accuracy (STARD) kontrol listesine bakılabilir (<https://www.equator-network.org/wp-content/uploads/2015/03/STARD-2015-checklist.pdf>).

Örnek: S-100B proteininin akut iskemik inmedeki tanısal doğruluğu araştırılsın. Acil servise inme semptomlarıyla gelen 200 hasta ardışık ele alınsın. Hastalarda S-100B proteini test edilsin. Testin eşik değeri **0.5 µg/L** olarak seçilsin. Toplam 200 hastadan 130’unda S-100B değeri >0.5 µg/L, 70’inde <0.5 µg/L olsun. Daha sonra referans yöntemlerle hastaların kesin tanısı konulsun. Son durumda S-100B’si eşik değerin üzerinde olan hastalardan (n=130) 90’ında iskemik inme saptansın, geri kalanında (n=40) iskemik inme tespit edilemesin. Buna karşılık, aynı semptomlara sahip ama S-100B değeri eşik değerin altında olan hastalardan (n=70) 10’unda iskemik inme saptansın, 60’ında saptanmasın. Bu verilere göre 2x2 tablosu düzenlensin:

	İnme (+)	İnme (-)
S-100B >0.5 µg/L	90 (DP)	40 (YP)
S-100B <0.5 µg/L	10 (YN)	60 (DN)

$$\text{Duyarlılık} = \frac{\text{DP}}{\text{DP} + \text{YN}} = 90/100 = \%90$$

Testin duyarlılığı %90'dır. Bu şu anlama gelir: Acil servise inme semptomlarıyla gelen ve S-100B değeri >0.5 µg/L olan hastaların %90'ında test sonucu doğrudur. Bu durumda test hastaların %10'unu kaçırmaktadır (YN).

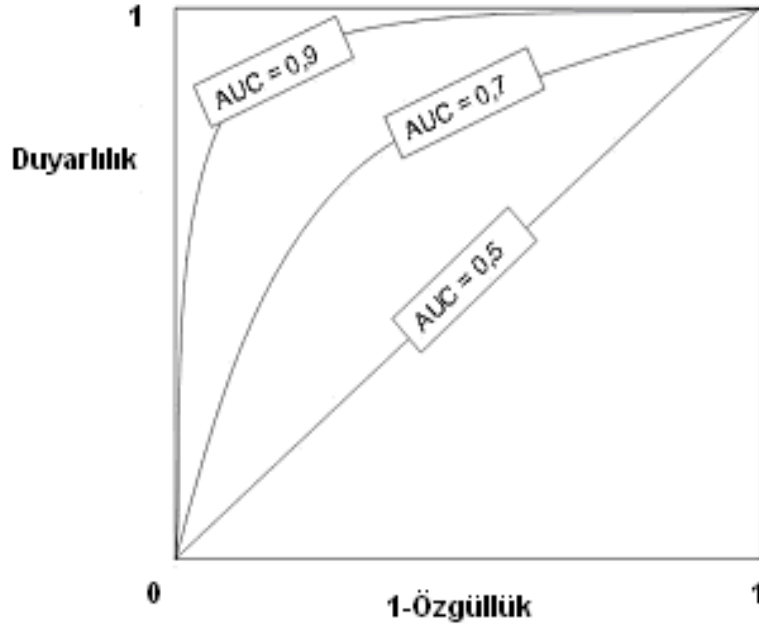
$$\text{Özgüllük} = \frac{\text{DN}}{\text{DN} + \text{YP}} = 60/100 = \%60$$

Testin özgüllüğü %60'tır. Bu şu anlama gelir, inme semptomlarıyla acil servise gelen hastalardan S-100B değeri <0.5 µg/L olanların %60'ında gerçekten inme yoktur. Ancak, referans yöntemle inme saptanmayan hastaların %40'ında test pozitif sonuç vermekte, S-100B değeri >0.5 µg/L eşiğini aşmaktadır. Bu şekilde acil servise gelen 10 hastadan 4'ünde test yanlış pozitif sonuç vermektedir.

Görüldüğü gibi, duyarlılık ve özgüllük testin hasta ve kontroller arasında ayırım gücünü vermektedir. Ancak sadece duyarlılık ve özgüllük klinisyenin kararı için yeterli değildir. Ayrıca duyarlılık ve özgüllük, yukarıda da belirtildiği gibi seçilen eşik değere göre birbirinin aleyhine değişir. Biri azalırken diğeri artar.

Receiver Operating Characteristics (ROC) eğrileri eşik değere göre duyarlılık ve özgüllüğün durumunu ortaya koyar ve testin klinik performansını değerlendirmek açısından çok daha güçlüdür (Şekil 3).

Genel olarak duyarlılık ve özgüllüğün toplamı >170 ise test yararlıdır. Klinisyenler bu eşiği iyi muayene ile aşabilirler. Buna göre duyarlılığı da özgüllüğü de %95 olan bir laboratuvar testi çok yararlıdır.



Şekil 3. ROC eğrileri ve eğri altında kalan alan tanısal testin ayırım gücünü verir. Düşey eksene testin duyarlılığı, yatay eksene ise testin 1-Özgüllük değeri işaretlenir. Eğri sol üst köşeye yaklaştığı ölçüde testin ayırım gücü artar. Eğri altında kalan alan (AUC) da bu bakımdan önemlidir. Bu değer 1'e yaklaştığı ölçüde testin ayırım gücü artar. AUC değerleri aynı amaçla kullanılan testlerin ayırım güçlerinin karşılaştırılması açısından da değerlidir. İdeal tanısal test için $AUC = 1$ 'dir (%100 duyarlılık, %100 özgüllük). Ancak, eğri altında kalan alan, %95 güven aralığı ile birlikte değerlendirilmelidir. Eğer $AUC \leq 0.5$ ise testin hiçbir değeri yoktur. AUC değerlerine göre testlerin kategorizasyonu aşağıdaki tabloda verilmektedir:

Tablo 1. AUC Değerlerine Göre Tanısal Doğruluk

AUC	Tanısal Doğruluk
0.9-1.0	En yüksek
0.8-0.9	Çok iyi
0.7-0.8	İyi
0.6-0.7	Yeterli
0.5-0.6	Yetersiz
<0.5	Yararsız

Öngörü Değerleri. Duyarlılık ve özgüllük pozitif ve negatif test sonuçlarının öngörü değeri hakkında bilgi vermez. Oysa, klinisyen pozitif test sonucu durumunda hastanın gerçekten hasta olma olasılığını veya negatif test sonucu durumunda kişinin hasta olmama olasılığını bilmek isteyecektir. Bu soruların yanıtı **pozitif ve negatif öngörü değerleri (PPV ve NPV)** ile verilebilir.

Pozitif Öngörü Değeri (PPV). Pozitif test sonucu alınan kişinin hasta olma olasılığıdır. Yüzde olarak verilir. Matematiksel olarak DP'lerin tüm pozitiflere (DP + YP) oranıdır. Testin +PPV'si %50 ise testi yapmanın bir anlamı yoktur.

$$\text{Pozitif prediktif değer} = \frac{DP}{DP + YP}$$

Negatif Öngörü Değeri (NPV). Negatif test sonucu alınan kişide hastalığın bulunmama olasılığıdır. Negatif test sonucu alınan sağlıklı kişilerin, başka deyişle DN'lerin tüm negatiflere oranıdır.

$$\text{Negatif prediktif değer} = \frac{DN}{DN + YN}$$

PPV ve NPV değerleri de 2x2 tablosundan hareketle hesaplanabilir. S-100B örneğine dönülecek olursa:

	İnme (+)	İnme (-)
S-100B >0.5 µg/L	90 (DP)	40 (YP)
S-100B <0.5 µg/L	10 (YN)	60 (DN)

Yukarıdaki PPV ve NPV formüllerini kullanarak:

$$PPV = 90/130 = \%69.2$$

$$NPV = 60/70 = \%85.7 \text{ bulunur.}$$

Bu sonuçlar şu anlama gelir:

PPV değerine göre, pozitif test sonucu elde edilen kişilerin %69'u gerçekten hastadır. Dolayısıyla PPV ne kadar yüksekse testin inandırıcılığı o derece yüksektir. Klinisyen burada S-100B değerini $>0.5 \mu\text{g/L}$ olarak gördüğünde akut iskemik inme olasılığının %69 olduğunu düşünecektir.

NPV değerine göre ise inme semptomlarıyla gelen ve negatif S-100B sonucu alınan 100 hastadan %86'sında hastalık (inme) gerçekten yoktur. Geri kalan kişiler ise gerçekten hastadır ancak YN sonuç elde edilmiştir. NPV değeri arttıkça negatif sonuçların inandırıcılığı artar. Klinisyene inme semptomlarıyla gelen hastada negatif S-100B sonucu verildiyse, hastada %86 olasılıkla inme olmadığını düşünecektir.

Öngörü Değerlerinin Prevalansa Göre Değişimi. Öngörü değerlerinin yorumunda en çok dikkat edilmesi gereken nokta, hastalık prevalansından etkilenmeleridir. Dolayısıyla testin öngörü değeri örneğin Acil Servis Nöroloji Yoğun Bakım Ünitesinde farklı, diğer kliniklerde farklı olacaktır. Çünkü inme prevalansı Nöroloji Ünitesinde çok daha yüksektir. Gerçekte bir testin aynı hastalık için duyarlılık ve özgüllüğü aynıdır ancak öngörü değeri prevalansa göre değişmektedir. Popülasyonda hastalık prevalansı arttığı ölçüde PPV artar. NPV ise hastalık prevalansı ile ters orantılı seyreder.

Örnek: S-100B örneğinde, referans teknikle kesin tanı konulduğunda toplam 200 hastadan 100'ünde inmenin doğrulandığı, geri kalan 100'ünde inmenin dışlandığı anlaşılıyor. Bu o acil servis için inme prevalansının %50 olduğunu gösterir. Demek ki, prevalans %50 olduğunda PPV %69, NPV %86'dır. Şimdi inme olmayan hasta sayısının 1000 olduğunu varsayalım. Bu durumda prevalans %9.1'e inecektir ($100/1100 = \%9.09$). Oysa duyarlılık ve özgüllük gene %90 ve %60'dır. Bu durumda 2x2 tablosu şöyle değişir:

	İnme (+)	İnme (-)
S-100B $>0.5 \mu\text{g/L}$	90 (DP)	400 (YP)
S-100B $<0.5 \mu\text{g/L}$	10 (YN)	600 (DN)

Yeni 2x2 tablosuna göre:

$$\text{PPV} = 90/490 = \%18.4$$

$$\text{NPV} = 600/610 = \%98$$

Bu sonuçlar şu anlama gelir: Pozitif test sonucu alınan 100 hastadan sadece 18'inde inme olabilir, geri kalan 72'sinde YP'lik vardır. Buna karşılık negatif test sonucu alınan 100 hastadan 98'inde inme yoktur.

Öngörü Değerleri ile İlgili Önemli Noktalar. Hastalık prevalansının düşük olduğu bir popülasyonda, testin duyarlılığı yüksek de olsa, YP oranı yüksek olacaktır, çünkü PPV düşüktür. Bu yüzden tümör belirteçleri malign hastalık prevalansının düşük olduğu asemptomatik popülasyonda iyi tarama testi ve kanser doğrulama testi değildirler.

Diğer önemli bir nokta, yukarıda da vurgulandığı gibi, PPV ve NPV genelleştirilemez. Örneğin birinci basamak sağlık hizmetlerinin sunulduğu sağlık kuruluşlarındaki prevalans ile üçüncü basamaktaki prevalans farklı olacaktır. Bu yüzden, örneğin BNP/ProBNP gibi testler üçüncü basamakta birinci basamağa göre daha güçlü klinik performans gösterecektir.

Olabilirlik Oranları (LR). LR'ler iki olasılığın birbirine oranıdır. Bu olasılıklardan birisi test sonucunun hastalarda görülme olasılığı, diğeri aynı test sonucunun sağlıklı kişilerde görülme olasılığıdır. Pozitif ve negatif LR'ler vardır (+LR ve -LR).

+LR, pozitif bir test sonucunun hasta kişide görülme olasılığının hasta olmayan kişide görülme olasılığına oranı anlamına gelir. Test sonucunun hastada görülme olasılığının, hasta olmayanda görülme olasılığından kaç kat yüksek olduğunun göstergesidir.

$$+LR = \text{Duyarlılık} / (1 - \text{Özgüllük}) = DP / YP$$

-LR ise bir negatif test sonucunun hasta bir kişide görülme olasılığının sağlıklı bir kişide görülme olasılığına oranıdır.

$$-LR = (1 - \text{Duyarlılık}) / \text{Özgüllük} = YN / DN$$

LR'lerin PPV ve NPV'ye üstünlüğü hastalık prevalansındaki değişikliklerden etkilenmemeleridir. Dolayısıyla farklı araştırmalardan elde edilen LR sonuçları başka bir popülasyona uygulanabilir. Ayrıca, LR'ler tanısal doğruluğun çok güvenilir göstergeleridir. Çünkü, test öncesi ve test sonrası olasılıklarla doğrudan bağlantılıdır. Klinisyen, LR'ler sayesinde baştaki düşüncesini test sonucuna göre değiştirebilir. Test öncesi ve test sonrası olasılıklar **Fagan Nomogramı** ile kolayca belirlenebilir.

Genel olarak +LR >10 ise test tanı koyma açısından çok değerlidir; -LR <0.1 ise tanıyı dışlama açısından çok değerlidir. +LR <10, -LR >0.1 ise testin katkısı nispeten azdır. LR = 0 ise, test hastalık olasılığına hiçbir katkıda bulunmaz.

Örnek. S-100B örneği için LR'leri hesaplayalım:

$$+LR = \text{Duyarlılık}/(1-\text{Özgüllük}) = DP/YP = 0.90/0.40 = 2.25$$

$$-LR = (1-\text{Duyarlılık})/\text{Özgüllük} = YN/DN = 0.10/0.60 = 0.16$$

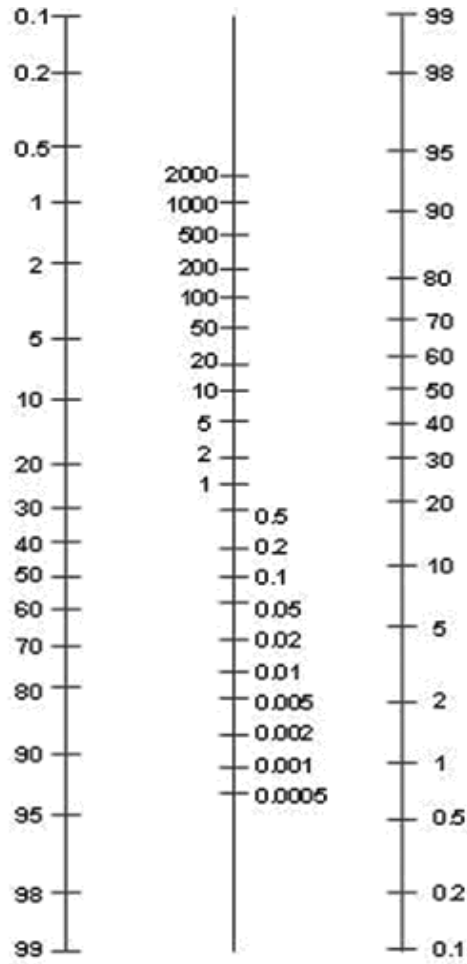
Tablo 2. Pozitif ve Negatif LR Değerlerinin Yorumlanması.

Testin Gücü	+LR	-LR
Çok iyi	>10	<0.1
İyi	6	0.2
Orta	2	0.5
Değersiz	1	1

Fagan Nomogramı. Klinisyen hasta hakkındaki kararını olasılıklarla verir. Hasta başvurduğunda klinisyen daha ilgili test isteminde bulunmadan bilgi ve deneyimlerine dayanarak, hastanın şikayetleri ve risk faktörleri ile hastalığın prevalansını sentezleyerek bir olasılık tahmininde bulunur. Bu test öncesi olasılığıdır (a priori probability veya pretest probability).

Bu bilgiden test sonrası olasılığa geçebilmek için LR'ye gerek vardır (Şekil 4).

Fagan Nomogramı, test öncesi olasılıktan (pratik olarak prevalans) LR değerini kullanarak test sonrası olasılığa ulaşılmasını sağlar.



Test öncesi olasılık Olabilirlik oranı Test sonrası olasılık

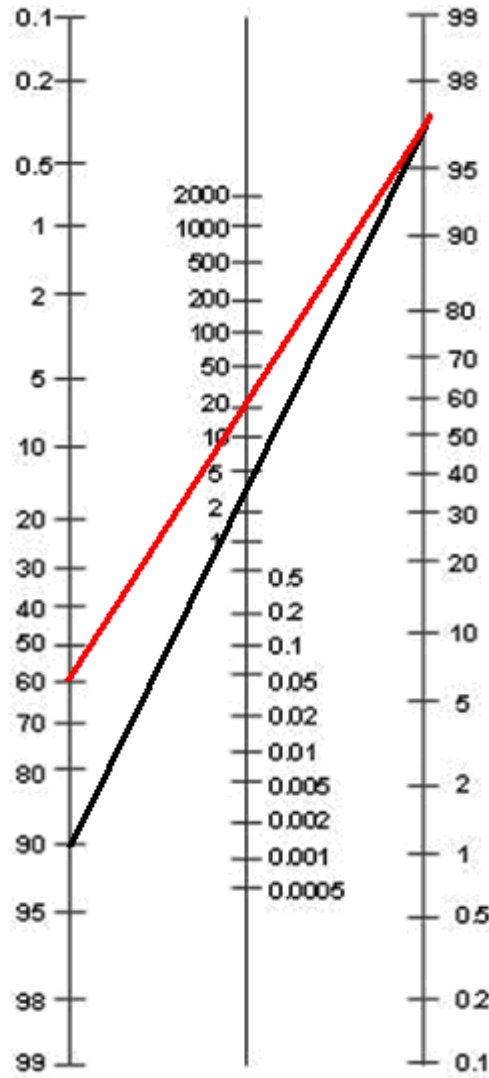
Şekil 4. Fagan Nomogramı. Test öncesi olasılık ve LR birleştirilerek test sonrası olasılık belirlenir. Eğer bir testin tanısal doğruluğu yüksekse, test sonrasında hastalık olasılığının belirlenmesine veya hastalığın dışlanmasına önemli katkıda bulunur. Eğer testin tanısal doğruluğu düşükse, test öncesi olasılıkta fazla bir değişiklik olmaz.

Acil servise gelen inme hastalarıyla devam edelim.

Acil servise 75 yaşında bir erkek hipertansiyon hastasının inme semptomlarıyla geldiğini varsayalım. Fizik muayene bulgularına, hastanın öyküsüne ve inme prevalansına dayanarak klinisyen hastada inme riskini %90 olarak düşünsün. Bu test öncesi olasılıktır. +LR = 2.25 olarak hesaplanmıştı. Fagan Nomogramı kullanılarak test sonrası olasılık hesaplanır. Bu değer yaklaşık %97'dir. Görüldüğü gibi, +LR düşük olduğundan test sonrası olasılıkta çok büyük bir değişim olmamaktadır (Şekil 5). Dolayısıyla S-100B çok da yararlı değildir.

Test öncesi olasılığın klinisyence %60 olarak saptandığını varsayalım. S-100B'nin +LR'si ise 20 olsun. Bu durumda test sonrası olasılık Fagan Nomogramı ile gene yaklaşık %97 olarak saptanır. Ancak, bu kez hastalık olasılığı %60'tan %97'ye yükselmiştir. Dolayısıyla S-100B tanıya varmada çok yararlıdır.

Buradan şu sonuca varmak mümkündür. Eğer hastada test öncesi hastalık olasılığı yüksekse, testin tanıya katkısı az olmaktadır. Eğer test öncesi olasılık düşükse ve testin tanısallık doğruluğu da yüksekse, test sonrası olasılık çok daha yüksek olmaktadır. Bu durumda test sonucu inme olasılığını önemli ölçüde değiştirecektir.



Test öncesi olasılık Olabilirlik oranı Test sonrası olasılık

Şekil 5. Fagan Nomogramı ile LR ve test öncesi olasılık değerleri kullanılarak test sonrası olasılığın hesaplanması. Görüldüğü gibi, test öncesi olasılık düşük olduğunda ve LR arttığında test sonrası olasılık artmaktadır. Test öncesi olasılık yüksekse, test sonrası olasılık üzerinde çok önemli katkı sağlaması için LR'nin çok yüksek olması gerekir.

ROC Eğrilerinin Önemi

ROC eğrileri, LR'lerin görülmesi açısından da değerlidir. Eğri üzerindeki herhangi bir nokta, seçilen eşik değer için +LR'yi verecektir. Dolayısıyla LR'ler seçilen eşik değere göre değişecektir. Örneğin Ferritin demir eksikliği anemisi için LR'leri seçilen eşik değerlere göre şöyle seyir göstermektedir (Tablo 3).

Tablo 3. Değişik Eşik Değerler İçin Ferritin Demir Eksikliği Anemisinde +LR'leri.

Ferritin (µg/L)	+LR
≥100	0.08
45 - 99	0.54
35 - 44	1.83
25 - 34	2.54
15 - 24	8.83
<15	51.85

LR'lerle ilgili diğer önemli bir nokta, testin özgüllüğü çok yüksek ise, duyarlılık düşük olsa bile +LR'nin yüksek olmasıdır.

Etkinlik (Efficiency). Testin doğru olarak saptadığı kesimin (tüm doğru pozitifler ve doğru negatifler) tüm çalışılan kişilere oranıdır.

$$\text{Testin etkinliği} = \frac{DP + DN}{DP + YP + DN + YN}$$

Tanımsal Doğruluk Ölçütleri İçin Kılavuz

Genel kural olarak genel popülasyonda gizli (belirtisiz) olan bir hastalığın (prevalansı düşük hastalıklar) taranması amacıyla kullanılan testlerin duyarlılığı mümkün olduğunca yüksek olmalı, yanlış pozitiflik oranı da (1-özgüllük) kabul edilebilir düzeyde olmalıdır. Ancak genel olarak istenen maksimum duyarlılık ve özgüllükte olmasıdır.

Tarama amaçlı kullanılan bir test için %95 duyarlılık ve %95 özgüllük durumunda, prevalans eğer %1 ise, pozitif test sonucu alınan hastaların %83.9'u hasta değildir (100-16.1). (Bu test için prevalans %1 olduğunda PPV %16.1 olmaktadır.) Eğer testin yanlış pozitiflik oranı yüksek bulunursa, bu durumda duyarlılıktan taviz verilerek özgüllük artırılabilir (kesim değeri değiştirilebilir). Ya da tersine, testin duyarlılığı ve pozitif prediktif değeri yetersiz bulunursa, test tarama amaçlı kullanılamaz, ama gene de prevalansı yüksek popülasyonlarda kullanılabilir. Veya aynı amaca yönelik bağımsız bir test de paralel çalışılarak PPV artırılabilir (Orthogonal Testing).

Örneğin bir testin duyarlılığı %90, özgüllüğü %95, diğer testin de duyarlılık ve özgüllüğü aynıysa (%90 ve %95), prevalansa göre PPV ortogonal test kullanımında aşağıdaki şekilde değişim gösterecektir.

Prevalans (%)	Tek Bir Testin PPV'si (%) (Duyarlılık %90, Özgüllük %95)	İki Test Birlikte PPV (%) (Duyarlılık %90, Özgüllük %95)
2	26.9	86.9
5	48.6	94.5
10	66.7	97.3
30	88.5	99.3

En yüksek klinik performans ne zaman gereklidir?

- **En yüksek duyarlılık (%100) şu durumlarda gereklidir:**
 - Hastalık ciddidir ve atlanmaması gerekir,
 - Hastalık tedavi edilebilir bir hastalıktır,
 - Yanlış pozitif sonuçlar hastada ciddi fiziksel, psikolojik ya da ekonomik travmaya neden olmamalıdır.

Örnek. Feokromositoma. Bu hastalık eğer atlanırsa fatal olabilir. Ama tanı konulursa neredeyse %100 tedavi şansı vardır.

Benzer şekilde fenilketonüri, venereal hastalıklar ve diğer tedavi edilebilir enfeksiyonlar için de aynı durum geçerlidir.

- **En yüksek özgüllük (%100) şu durumlarda gereklidir:**

- Hastalık ciddidir ama aynı zamanda tedavi edilemezdir,
- Hastalığın olmadığıнын bilinmesi psikolojik olarak ya da halk sağlığı açısından önemli değildir,
- Yanlış pozitif sonuçlar hastada ciddi psikolojik ve ekonomik travmaya yol açar.

Örnek. Multiple skleroz ve çoğu gizli kanser durumlarında. Bu hastalıklar ciddidir, ancak genellikle tam anlamıyla tedavi edilemezler.

- **Yüksek pozitif prediktif değer şu durumlarda gereklidir:**

- Yanlış pozitif test sonucu alınan hastanın tedavisi ciddi istenmeyen sonuçlara yol açabilir.

Örnek. Gizli akciğer kanseri. Bu durumda lobektomi ya da radyasyon tedavisi uygulanabilir ki sonuçları fatal olabilir. Veya yanlış pozitif beta-HCG sonuçlarının gereksiz pek çok ameliyata neden olduğu bilinmektedir.

- **En yüksek etkinliğin gerekli olduğu durumlar:**

- Hastalık ciddidir ama tedavi edilebilir,
- Gerek yanlış pozitif, gerekse yanlış negatif sonuçlar eşit ölçüde ciddi sorunlar yaratır.

Örnek. Miyokart infarktüsü. Hastalık fataldir ama tedavi edilebilir. Lupus eritematosus, bazı lösemi ve lenfoma türleri, diabetes mellitus.

Odds Oranı ve Bayes Teoremi

Odds Oranı. Belli bir hastalığın var olma olasılığının olmama olasılığına oranıdır. Hastalığın popülasyondaki prevalansını yansıtır.

Örnek. Yetmiş beş yaşında bir erkekte 1.3 cm^3 büyüklüğünde prostat karsinoması görülme olasılığı yaklaşık olarak %8'dir. Buna göre, 70 yaşından büyük bir erkekten otopsi sonrasında yapılan prostat seksiyonunda 1.3 cm^3 'den büyük histolojik prostat Ca bulgusu için odds oranı

$$0.08 / (1-0.08) = 1:11.5 \text{ olur.}$$

Bayes Teoremi. Bayes teoremi, hastalık ile ilgili eski bilgilere yeni bilginin eklenmesinden sonra hastalık olasılığını hesaplamakta kullanılan bir yöntemdir. Kullanılan formüller, PV hesabındaki gibidir. Prevalans yerine, test yapılmadan önceki bilgilere dayanan olasılık kullanılır. Bu formüller, bilgisayar programları ile uygulanmaktadır. Bilgisayarsız olarak yapılabilen bir yöntem de vardır.

Burada önce test sonucu bilinmeden önce hastalığın gerçekleşmesi ile ilgili odds oranı hesaplanır. Daha sonra bu oran, test sonucu ile olabilirlik oranı şeklinde kombine edilir. En sonunda elde edilen veri gene odds oranı şeklindedir ve istenirse olasılığa dönüştürülebilir. Bu yöntemin avantajı, kolay akılda kalması ve az matematiksel işlem gerektirmesidir:

$$\text{Odds oranı (sonra)} = \text{odds oranı (önce)} \times \text{olabilirlik oranı}$$

Örnek: PSA değeri 4.0-10 ug/L arasında olan bir benign prostat hipertrofisi (BPH) hastası için ürolog transrektal ultrasonografi (US) istiyor ve pozitif sonuç alıyor. Ancak, eski deneyimleriyle biliyor ki, her iki testin de pozitifliği sonrasında biyopsi yaptığında negatif sonuç alabiliyor. Dolayısıyla her iki tarama testinin de yanlış pozitifliği olduğunu biliyor. Ve bu hastada Prostat Ca olasılığını hesaplamak istiyor.

Bu durumda önce PV'ye bakılabilir. Buna göre $\text{PSA} > 4.0 \text{ ug/L}$ olduğundan, bu hastada Prostat Ca olasılığı %17'dir diyebilir. Ancak, %17, tüm bu gruba giren ($\text{PSA} > 4.0 \text{ ug/L}$ olan hastalar) içindir. Bunların içinde bazı çok yüksek PSA değeri olanlar da vardır. Bu BPH hastalarında ($\text{PSA} 4.0-10.0 \text{ ug/L}$ olanlarda) biyopsi sonucunda Prostat Ca tanısı %12 kadardır.

Bu durumda:

1. US yapılmadan önce karsinoma varlığı için odds oranı nedir?

Biyopsi sonucunda bu grupta karsinoma görülme olasılığı %12 idi. Hastalığın olmama olasılığı ise $1.0 - 0.12 = 0.88$ olur. Bu durumda odds oranı:

$$0.12/0.88 = 0.14 \text{ bulunur.}$$

2. Transrektal US bulgusunun olabilirlik oranı nedir?

Eski bilgilere göre duyarlılık yaklaşık %92, özgüllük ise %30-%70 arasındadır (ortalama %50). Bu durumda pozitif olasılık oranı:

$$\text{Duyarlılık} / (1 - \text{özgüllük}) = 0.92 / (1 - 0.50) = 1.8 \text{ bulunur.}$$

3. Yeni bilgi eklendikten sonraki kombine odds oranı ne olur?

Odds oranı (sonra) = odds oranı (önce) x olabilirlik oranı idi

$$\text{Odds oranı (sonra)} = 0.14 \times 1.8 = 0.252$$

4. Odds oranı olasılığa nasıl dönüştürülür?

Olasılık = odds oranı / (1 + odds oranı) şeklinde hesaplanır.

$$\text{Olasılık} = 0.25 / (1 + 0.25) = 0.25 / 1.25 = 0.2 \text{ bulunur.}$$

Buna göre, her iki test de pozitif olsa biyopside pozitif sonuç olasılığı %20'dir.

Örnek: D-Dimer testi pozitif bir hastada akciğer embolisi olasılığı nedir? Test öncesi olasılığının %28 olduğu düşünülmektedir. D-dimer için duyarlılık %90, özgüllük %95'dir.

1. Test öncesi olasılık odds oranına dönüştürülür:

$$\text{Odds oranı} = \text{Olasılık} / (1 - \text{Olasılık}) = 0.28 / (1 - 0.28) = 0.389$$

2. +LR hesaplanır:

$$+LR = \text{Duyarlılık} / (1 - \text{Özgüllük}) = 0.90 / (1 - 0.95) = 18$$

3. Test sonrası odds oranı hesaplanır:

$$\text{Test sonrası odds oranı} = \text{Test öncesi odds oranı} \times +LR = 0.389 \times 18 = 7.0$$

4. Test sonrası odds oranı olasılığa dönüştürülür:

$$\text{Test sonrası olasılık} = (\text{Test sonrası odds oranı}) / (1 + \text{Test sonrası odds oranı}) = 7 / (1 + 7) = 7/8 = \%87.5$$

Tablo 4. Özet Olarak Tanısal Doğruluk Ölçütleri.

Ölçüt	Hesaplama	Açıklama
Doğru pozitif (DP)	2x2 tablosundan alınır	Pozitif test sonucu olan hastalar
Doğru negatif (DN)	2x2 tablosundan alınır	Negatif test sonucuna sahip hasta olmayan kişiler
Yanlış pozitif (YP)	2x2 tablosundan alınır	Test sonucu pozitif olan kişilerde hastalığın olmaması
Yanlış negatif (YN)	2x2 tablosundan alınır	Negatif test sonucu olan hastalar
Duyarlılık = Doğru pozitiflik oranı	$DP / (DP + YN)$	Hastada pozitif test sonucu olasılığı
1 - Duyarlılık = Yanlış negatiflik oranı	$YN / (DP + YN)$	Hastada negative test sonucu olasılığı
Özgüllük = Doğru negatiflik oranı	$DN / (DN + YP)$	Hasta olmayan bir kişide negative test sonucu
1 - Özgüllük = Yanlış pozitiflik oranı	$YP / (DN + YP)$	Hasta olmayan kişide pozitif test sonucu olasılığı
PPV	$DP / (DP + YP)$	Pozitif test sonucu alınan kişinin hasta olma olasılığı
NPV	$DN / (DN + YN)$	Negatif test sonucu alınan bir kişinin hasta olmama olasılığı
Doğruluk/Etkinlik	$(DP + DN) / (DP + DN + YP + YN)$	Test sonuçlarının hastalık varlığı veya yokluğunu doğru saptaması
Bayes teoremi	Test sonrası Odds = Test öncesi Odds X LR	Testten sonra hastalığın bulunma olasılığının bulunmama olasılığına oranı
+LR	Duyarlılık / (1 - Özgüllük)	Pozitif test sonucu sonrasında hastalığın odds oranında artış
-LR	(1 - Duyarlılık) / Özgüllük	Negatif test sonucu durumunda kişinin hasta olma olasılığında azalma

Takipte Laboratuvar Testleri, Biyolojik Değişkenlik ve Referans Değişim Değeri

Başta da belirtildiği gibi, laboratuvar testlerinin başlıca istem nedeni tarama/risk belirleme, tanı ve tedavi takibidir. Tedavi takibi veya hastalığın izlenmesinde test tekrarı gereklidir. Örneğin diyabette HbA1c ve glukoz takibi; gastrointestinal kanaması olan hastada hematokrit takibi; böbrek fonksiyonu bozulan hastada kreatinin takibi gibi.

Takipte kullanılan testlerin hem analitik performansı, hem de analitin bireysel biyolojik değişkenliği çok önemlidir.

Her bir analitin kişiye özgü bir değişkenliği vardır. Buna **bireysel biyolojik değişkenlik** denilir. Hasta takibinde bu değişkenlik göz önüne alınmalıdır. Analitin bireyselliği referans aralıkların kullanılması hakkında da önemli bilgi sağlar. Bu amaçla **Bireysellik İndeksi** kavramı geliştirilmiştir. Bireysellik indeksi pratik olarak **bireysel biyolojik değişkenliğin (CVi) bireyler arası biyolojik değişkenliğe (CVg) oranıdır:**

$$\text{Bireysellik İndeksi (Bİ)} = \frac{\text{CVi}}{\text{CVg}}$$

CVi <<CVg ise, bu testin bireyselliğinin yüksek olduğunu gösterir ve bu durumda Bİ düşüktür. Tersine, Bİ ne kadar yüksekse testin bireyselliği o derece düşüktür. Bireyselliği yüksek testler için popülasyona dayalı referans aralıklar çok değer taşımaz. Referans aralıklar, bireyselliği düşük testler (Bİ değeri yüksek) için yararlıdır. Öte yandan, bireyselliği yüksek testler takip açısından çok yararlıdırlar.

Genel olarak Bİ <0.6 ise bu testler için referans aralıklar uygun değildir (örneğin kreatinin, Bİ = 0.33). Çünkü kişiye bağlı değişim referans aralık içinde dar bir bölgede olur ve referans aralıklarla karşılaştırarak kişideki değişimi saptamak mümkün olmaz. Bu testler bu yüzden takipte çok yararlıdır. Bİ >1.4 olan testler ise bireyselliği çok düşük testler olarak görülür (örneğin serum demiri, Bİ = 1.14) ve bu testler için referans aralıklar kullanılarak test sonuçlarından tanı ve taramada yararlanılır. Yaşa, cinsiyete vb. alt gruplamalarla CVg'yi düşürerek Bİ'yi yükseltmek ve böylece referans aralıkları sağlıklı kullanmak mümkün olabilir.

Referans Değişim Değeri. Takipte referans değerlerin veya sabit değerlerin kullanımı yerine daha dinamik ve fonksiyonel **Referans Değişim Değeri (RDD)** kavramı geliştirilmiştir. RDD, analitin **hem biyolojik değişkenliğini, hem de analitik değişkenliğini** esas alır. Belli olasılık düzeylerinde kişinin sonuçları arasında önemli fark olup olmadığının saptanmasında kolaylık sağlar. RDD şu formülle hesaplanmaktadır:

$$\text{RDD} = 2^{1/2} * Z * [\text{CVa}^2 + \text{CVi}^2]^{1/2}$$

Burada **CVa** analitik deęişkenlik; **CVi** bireysel biyolojik deęişkenliktir. Z deęeri %95 olasılık düzeyinde tek yönlü düşünülürse 1.65, çift yönlü düşünülürse 1.96; %99 olasılık düzeyinde tek yönlü düşünülürse 2.33, çift yönlü 2.58'dir. $2^{1/2}$ deęeri, iki ölçüm arasındaki fark araştırıldığından formüle girmiştir.

Örnek: Kolesterol yükseklięi için statin tedavisi uygulanan bir hastada ilaca baęlı karacięer enzimlerindeki yükselmenin önemli olup olmadığı araştırılsın. Hastanın ALT ve AST deęerleri zaman içinde ařağıdaki tabloda verildięi gibi seyretsin. ALT ve AST için analitik deęişkenlik $CVa = \%10$ olsun. Buna göre ALT ve AST için RDD hesaplanacak olursa:

ZAMAN	ALT (U/L)	AST (U/L)
Başvuru	30	28
İlaç öncesi	33	30
2 ay sonra	45	40
4 ay sonra	60	62
6 ay sonra	80	88

ALT

%95 (Z=1.96); $CVa = \%10$ ve $CVi = \%15.6$ için:

$$RDD = 1.414 * 1.96 * [10^2 + 15.6^2]^{1/2} = \%51.4$$

%99 (Z=2.58) ve $CVa = \%10$ için

$$RDD = 1.414 * 2.58 * [10^2 + 15.6^2]^{1/2} = \%67.6$$

AST

$CVa = \%10$ ve $CVi = \%4.7$ için

%95: %30.6

%99: %40.5

Bu sonuçlar ilaç kullanımından 4 ay sonra ALT ve AST deęerlerinin bazal deęere göre %100 oranında arttıęını, dolayısıyla %99 olasılık düzeyindeki farkın bile ařıldığına, sonuç olarak karacięerin statinden olumsuz etkilendięini göstermektedir.

KAYNAKLAR

1. Shultz EK, Aliferis C, Aronsky D. Clinical evaluation of methods. In: Burtis CA, Ashwood ER, Bruns DE (Editors). Tietz textbook of clinical chemistry and laboratory medicine. 4th edition. St. Louis: Elsevier Saunders; 2006, pp: 409 – 424.
2. Solberg HE. Establishment and use of reference values. In: Burtis CA, Ashwood ER, Bruns DE (Editors). Tietz textbook of clinical chemistry and laboratory medicine. 4th edition. St. Louis: Elsevier Saunders; 2006, pp: 425 – 448.
3. Young DS, Bermes EW. Preanalytical variables and biological variation. In: Burtis CA, Ashwood ER, Bruns DE (Editors). Tietz textbook of clinical chemistry and laboratory medicine. 4th edition. St. Louis: Elsevier Saunders; 2006, pp: 449 – 473.
4. Kaplan C. Use of the laboratory. In: Walker HK, Hall WD, Hurst JW (Editors). Clinical methods: The history, physical, and laboratory examinations. 3rd edition. Boston: Butterworths; 1990. pp: 40-48.
5. Simon D, Boring JR. Sensitivity, specificity, and predictive value. In: Walker HK, Hall WD, Hurst JW (Editors). Clinical methods: The history, physical, and laboratory examinations. 3rd edition. Boston: Butterworths; 1990. pp: 49-54.
6. Fraser CG. Biological variation. From principles to practice. Washington: AACC Pres; 2001.
7. Wians FH. Clinical laboratory tests: Which, why, and what do the results mean? Labmedicine 2009;40:105-113.
8. Higgins C. An introduction to reference intervals (1) - some theoretical considerations. www.bloodgas.org, 2009.
9. Horvath AR, Kis E, Dobos E. Guidelines for the use of biomarkers: Principles, process and practical considerations. Scand J Clin Lab Invest 2010;70(Suppl 242): 109-116.
10. Wick MR, Marchevsky AM. Evidence-based principles in pathology. Existing problem areas and the development of “quality” practice patterns. Arch Pathol Lab Med 2011;135:1398-1404.
11. Koenig W. Integrating biomarkers: the new frontier? Scand J Clin Lab Invest 2010;70(Suppl 242): 117-123.

12. Ceriotti F, Hinzmann R, Panteghini M. Reference intervals. *Ann Clin Biochem* 2009;46: 8-17
13. Fraser CG. Inherent biological variation and reference values. *Clin Chem Lab Med* 2004;42:758-764.
14. Islin H. Interpretation of laboratory results. www.bloodgas.org, 2010.
15. Simundic A-M. Diagnostic accuracy part 1. Basic concepts: sensitivity and specificity, ROC analysis, STARD statement. www.bloodgas.org, 2001.
16. Simundic A-M. Diagnostic accuracy – part 2. Predictive value and likelihood ratio. www.bloodgas.org, 2009.
17. Pickup JF, Haris EK, Kearns M, Brown SS: Intra-individual variation of some constituents and its relevance to population-based reference ranges. *Clin Chem* 1977;23:842-850.
18. Harris EK, Brown SS. Temporal changes in the concentrations of serum constituents in healthy men. Distributions of within-person variances and their relevance to the interpretation of differences between successive measurements. *Ann Clin Biochem* 1979;16:169-176.
19. EFLM biological variation database. <https://biologicalvariation.eu/>, 2019.