

TIPDA ÖLÇME VE DEĞERLENDİRME HATALARI

Dr. M. Rahmi Dirican(x)

1. Giriş:

Tıp ilmi ilerledikçe hastalıkların tanısı ve tedavisiyle ilgili yöntemler değişmekte ve daha iyiye doğru gitmektedir. Bir yüzyıl öncesine kadar, hemen sadece fizik muayeneden ibaret olan hastalık tanısı yöntemlerine her geçen gün yenileri eklenmektedir. Bunlar arasında-fizik muayene önemini kaybetmemiş olmakla beraber-laboratuvar muayenelerinin çok önemli bir yeri vardır. Böylece daha kesin ve daha güvenilir bir şekilde sonuca varılmaktadır. Tedavi yöntemleri de değişmiştir. Yakın zamanlara kadar, tanınmış bir hekimin «kanaatimce, A hastalığının tedavisinde X tedavisi bütün diğer tedavilerden iyidir» demesi o hastalığın tedavisini derhal değiştirmeye yetmekteydi. Bugünün hekimi, tanınmış bir hekimin kişisel görüşünü derhal benimsemekten çok, inandırıcı ve objektif veriler görmek istemektedir. Ve araştırmalarla sağlanabilen inandırıcı veriler arasında kişisel görüşe yer yoktur.

Hastalıkların tanısında ve tedavisinde laboratuvar yöntemleri önem kazandıkça ve bunların temeli de ölçme ve bu ölçmeyi değerlendirme oldukça,

ölçme ve değerlendirmedeki hataların tesbit edilmesi gerekli olmuştur. Bilimsel çalışan bir hekim, çalışmalarından sonuç çıkarırken bir istatistikçinin yardımını isteyebilir ve ona «ben şu kadar veri topladım, bu verileri değerlendirelim ve bir sonuca varalım» diyebilir. Hekim, örneğin, prostatektomi yapılan hastalarda ameliyat öncesi ve sonrası kan üresi miktarıyla ameliyat sonu ölümler arasında bir ilişki olup olmadığını araştırmaktadır. Hekim, bütün muayene sonuçlarını toplamış ve sadece değerlendirme için istatistikçinin yardımını istemiş olabilir. Bu gibi durumlarda istatistikçinin ilk yapacağı iş, mevcut verilerin güvenilir olup olmadığını araştırmak olacaktır. Çünkü güvenilir olmayan verilere dayanarak doğru bir sonuca varılamaz. İstatistikçi araştırmayla ilgili verilerin güvenilir olup olmadığını incelerken, araştırmacının bilimsel gücünden ya da doğruluğundan şüphe ettiği için değil, araştırmanın planlanmasında bilimsel yöntemlere uyulup uyulmadığını yada ölçme hataları olup olmadığını, tesbite çalışır. Belki, örneğin, ameliyat öncesi ve ameliyat sonrası kanda üre miktarı ile ameliyat sonu ölümlerin ilgisini araştıran hekim, hastaların diğer niteliklerinin benzer olmasına yeteri kadar önem vermemiş

ve böylece kanda üre miktarı az olanların daha genç yaş grubunda toplandığının farkına varmamış olabilir. Bu yüzden, araştırmasının sonunda, kanda üresi az olanlarda ameliyat sonu ölümlerin az olduğu sonucuna varabilir. Ne varki böyle bir durumda, «hastaların daha genç oluşu mu, yoksa kanda üre seviyesinin nisbeten daha düşük oluşu mu ameliyat sonu ölümlerin azlığında rol oynamıştır?» sorusuna inandırıcı bir cevap verilemez. En iyisi her türlü araştırmanın bilimsel yöntemlere göre planlanması ve istatistikçiden sadece elde edilen sonuçların değerlendirilmesinde değil, araştırmanın planlanmasında da yardım istenmesidir.

Diğer bir ihtimal, araştırma devam ederken, laboratuvarda çalışan personelin ya da tahlil metodunun değişmesidir. Hatta, tahliller aynı metodla ve aynı personel tarafından yapılmış olsa bile laboratuvar teknisyeninin günlük yaşantısıyla ilgili çeşitli durumlar bulguları etkileyebilir. Araştırma yapan, kendisinin ve kendisiyle birlikte çalışan personelin günlük yaşantılarıyla ilgili olayları kaydettiği bir defterle araştırma sonuçlarını kıyaslayabilse laborant S. nin dişinin ağrıdığı gün veya laborant N. nin çocuğunun hastaneye kaldırıldığı gün kontrollü tahlil sonuçlarında büyük farklar olduğunu görecektir.

2. Ölçme ve Değerlendirme Hataları:

Tıpta, hemen bütün araştırmalar ve ölçmeler canlılar üzerinde yapıldığından, araştırma sonuçlarının tam manasıyla objektif olamayacağını ve üzerinde araştırma yapılan canlılarda meydana gelen değişikliklerin, araştırma sonucunu etkileyebileceğini önceden kabul etmek gerekir. Araştırmacı,

topladığı verileri mümkün olduğu kadar objektif hale getirmeğe çalıştıkça ve gözetleyicinin (ölçenin) değerlendirmede yapabileceği hataları gözününe aldıkça bilimsel hareket eder. Gerçekten, gözetleyicinin değerlendirmede değişik kararlara varması oldukça sık görülür. Bu hususta bazı örnekler vermek isteriz.

a- Londra'da bir araştırmacı sitratla muamele edilmiş aynı kandan örnekler alarak, bir saat içinde muayene edilmesi ricasıyla, Londra'nın en güvenilen 21 laboratuvarına göndermiştir. Çeşitli laboratuvarların bulguları Şekil:1 de görülmektedir. Şekilde, aranan her madde için bu laboratuvarların buldukları değerlerin aritmetik ortalaması 100 olarak kabul edilmiş ve her laboratuvarın bulgusunun bu ortalamadan ayrılış yüzdesi hesaplanarak bir yuvarlakla gösterilmiştir (1).

b- Şekil:2 de bir laboratuvarda özel olarak hazırlanmış bir eriyikten alınan örneklerin, Londra'nın en güvenilir 36 laboratuvarında aynı anda yapılan tahlillerinin sonuçları görülmektedir. Hazırlanan eriyikte bulunan çeşitli maddelerin teorik değerleri 100 kabul edilerek, çeşitli laboratuvarların bulgularının bu teorik değerlerden sapması yuvarlaklar halinde gösterilmiştir(1)

c- Mikrofilmleri değerlendiren 6 okuyucu birbirlerinden tamamen bağımsız olarak 1807 kişiye ait mikrofilmleri incelenmişlerdir. Değerlendirenlerin birisi bu mikrofilmlere göre 38, diğeri 83 kişiye tüberküloz teşhisi koymuştur. Tek okuyucunun bulguları gözününe alınırsa, tüberkülozlu olan hastaların %32 sinin tek okuyucu tarafın-

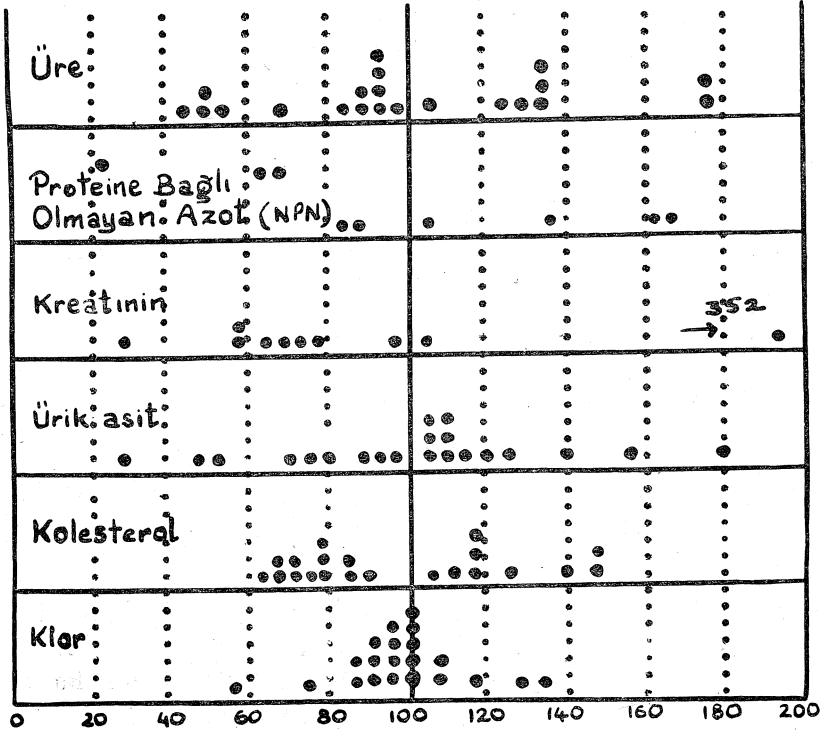
dan teşhis edilemediği ve tamamen sağlam olduğu bilinen kişilerden % 2 sine tüberküloz teşhisi konduğu ortaya çıkmaktadır (2).

d- Son olarak vereceğimiz örnek ilginçtir. New York'da 150 hastanın 30x40 ebadında ve aynı anda çekilen çift filmleri, üç radyoloji ve üç göğüs hastalıkları profesörüne okutturulmuştur. Aradan 3 ay geçtikten sonra aynı filmler, aynı şahıslara tekrar okuturulmuştur. İki profesörün aynı fikirde

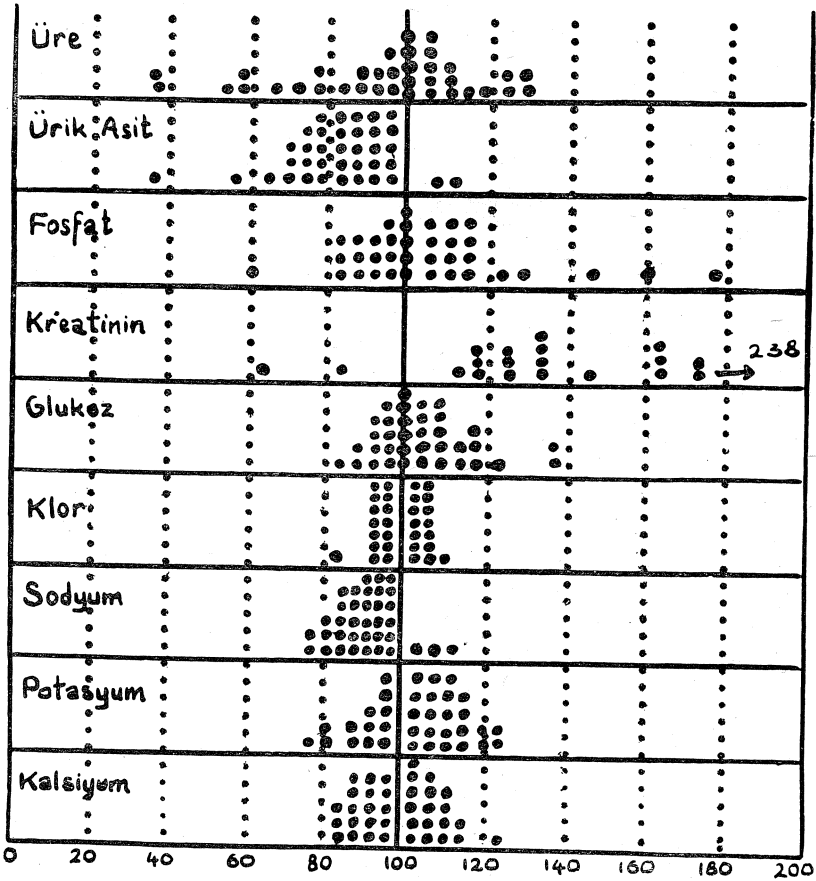
olmayışı, yaklaşık olarak vak'aların üçte birinde (%33) tesbit edilmiş ve aynı profesörün birinci ve ikinci defa aynı film hakkındaki kanaat uyuşmazlığı ise %10 oranında bulunmuştur(3).

3. Hata nedenleri:

Ölçme ve değerlendirme hatalarının nedenlerin çok çeşitlidir. Bunların en başında kriter yönünden anlaşmazlıklar gelmektedir. Bugün her hekimin kabul ettiği bir kriter vardır. Değişik fakültele-



Şekil: 1- Bir kan numunesinin değişik laboratuvarlar tarafından yapılan tahlil sonuçları (Her laboratuvarın bulgusu bir yuvarlakla gösterilmiştir. Bütün laboratuvarların bir madde için buldukları değerlerin ortalaması 100 kabul edilmiş ve her laboratuvarın bulgusunun bu ortalamadan ayrılış yüzdesi hesaplanarak bir yuvarlakla belirtilmiştir).



Şekil: 2 Özel olarak hazırlanmış bir eriyiğin 26 laboratuvar tarafından yapılan tahlil sonuçları (sonuçlar, hazırlanan eriyiğin teorik değerlerin yüzdesi olarak belirtilmiştir).

lerden mezun hekimler arasında kriter anlaşmazlıkları daha fazladır.

Bir an için, kriter anlaşmazlığı olmadığını kabul ederek, kolay ve basit gibi görünen bir ölçmeyi, arteriyel tansiyon ölçülmesini, ele alarak ölçme ve değerlendirme hatalarının nedenlerini inceleyelim. Hergün aynı hastaların tansiyonunu, birkaç dakika ara ile ile ölçen iki hekimin bulguları arasında

mutlaka fark olacaktır. Sistolik ve diastolik basınç kriterleri aynı olsa bile bu fark görülecektir. Bu farkın nedenleri şunlardır:

a- Gerçek Fark: Buna biyolojik değişiklik de denilebilir. Belirli bir anda ölçülen arter basıncı hiçbir zaman aynı seviyede kalmaz. Yapılan araştırmalar, çeşitli faktörlerin etkisiyle 24 saat içinde sistolik basınçta 90 mm.,

diastolik basınçta 40 mm. bir değişme olabileceğini göstermektedir (4). Bu değişimin nedenleri arasında günlük faaliyetlerin ve kişideki duygusal değişikliklerin (heyecan, üzüntü vb.) büyük rolü vardır. Sıcaklık ve mevsim değişikliklerinin kan basıncını etkilediği bilinmektedir. Nitekim, yazın kan basıncı yüksek olabileceği gibi, sıcak bir yerde ölçülen kan basıncının, soğuk yerde ölçülenden daha değişik sonuç verdiği bilinmektedir (5). Ayrıca, kan basıncı ölçülürken hastanın yatar ya da oturur vaziyette oluşu sonucu etkilemektedir (6).

b- Ölçme Farkı: Bu fark kullanılan farklı tansiyon aletlerinden veya farklı büyüklükteki manşetlerden veya manşetin kolda farklı yerlere konulmasından olabileceği gibi, alete ait hatalardan da ileri gelebilir. Manşetin eninin önemi de büyüktür. Alete ait şu hataları da sayabiliriz: Belki cıva basıncını belirten 0, 20, 120 vb. gibi rakamlar tam yerine konmamıştır veya cıvanın içinde yükselip alçaldığı cam boru kirlidir ve camın kirli yüzeyi cıvayı tuttuğu için sonucu etkilemektedir.

Ölçme farkı gözetleyicinin (ölçenin) yaptığı hatalardan da olabilir. Bu hataları iki grupta incelemek mümkündür:

(1) Sistemik hatalar: Bazı hekimler tansiyonu daima yüksek, bazıları ise daima biraz düşük okur. Her iki halde de ortalama değişir. İşte bulguların ortalama değerine tesir eden bu cins hatalara sistemik hatalar denilir. Bu hatalar ölçmenin tekniği, yeteneği ve değerlendirme kriteriyle ilgilidir. Bazı insanlar tansiyon ölçerken lastik pompaya hızlı basar veya havayı

çabuk boşaltır, bazıları bu işi daha yavaş yapar. Bu ölçme tekniği, sonucu etkiler. Kriter olarak bir hava verişte veya boşaltışta 0,5 cm. cıva seviyesini almak ve buna alışmak iyi olur. Tansiyon ölçenin konsantrasyon yeteneği yani sesin ilk duyuluşuna veya kayboluşuna hemen uyum yapabilmesi, işitme organının tam veya hafif arızalı oluşunun etkisi, ses işitilince veya kaybolunca görme organının cıva sütununu okuma hızı, veya gördüğü (okuduğu) değeri hafızasında tutup sonunda hatasız olarak kaydedebilmesi gibi hususlar sayılabilir. Bu sonuncu hata, tahmin edildiğinden çok daha fazla olur. Bazıları sadece cıva sütunu oynarken, bazıları da cıva sütununun hareketiyle beraber sesin duyulduğu zamanki değeri kaydederler. Bu gibi farklı değerlendirmeler de hatalı ölçmelere yol açabilir.

(2) Sistemik olmayan hatalar: Bunlar öyle hatalardır ki, iki okuyucunun bulgularının ortalamasına tesir etmezler, fakat yine de hatadır. Örneğin, okuyucu A ölçerken değerleri daima sonu 0 veya 5 olan milimetrelerle, B ise 2 ve 7 ile biten milimetrelerle belirtebilir. Bu iş, çok zaman okuyucunun farkına varmadığı bir husustur. Sistemik olmayan hatalardan bir diğeri de okuyan şahsın, tansiyon bulgusunu, çoğu zaman *peşin hükümlerin etkisi altında* değerlendirmesidir. Örneğin, hastanın anamnezi ve diğer fizik muayene bulguları, hekimde hipertansiyon kanısı yaratmışsa,

o hekim tansiyonu okurken biraz yüksek okumak eğilimindedir.

Bütün bu hataların yanı sıra, aynı hekim tarafından aynı hastada okunan değerler, iki ölçme arasında hastada vukubulacak gerçek farktan dolayı, değişik olabileceği gibi; bizzat tansiyonu ölçen hekimde de bu gerçek fark ne niyle değişiklik olabilir. Yani ölçmeyi yapan kişi, her zaman aynı dikkat, ilgi ve psişik durumu gösteremeyeceğinden ölçme farkları meydana gelebilir. Bu bakımdan iyi bir araştırmacı daima «ben kendi kendimle ne dereceye kadar uyuşmaktayım?» sorusunu sormalıdır.

4. Sonuç:

Ölçme hataları toplum sağlığı yönünden çok önemlidir. Kişiyi sağlıklı kılmayı, yani onu bedenlen, ruhen ve sosyal bakımdan daha iyi duruma getirmeyi amaç edinen bir hizmette çalışanların ölçmelerde yapılacak bu hataları gözönüne alarak teşhis ve değerlendirmeye çok dikkat etmeleri gerekir. Hastaların kapı önüne yığılmış olduğu bir poliklinikde, alelacele ölçülen ve yüksek bulunan bir tansiyonla, o hastaya hipertansiyon teşhisini koymadan önce, böyle bir teşhisin kişinin günlük yaşantısını tamamen değiştirebileceği ve bu değişikliğin o kişinin bağlı olduğu ailenin diğer bireylerinde ve dolayısıyla toplumda yaratabileceği sorunları daima gözönüne almak gerekir. Hasta-hekim ilişkilerinde bu hususa dikkat etmek iyi bir hekimin niteliklerindendir.

Yararlanılan Kaynaklar

1. Wootton, I.D.P., and King, E.J.: Normal values for blood constituents (inter-hospital differences), *Lancet*, 1: 470, 1953.
2. Yerushalmy, J. at all: The role of dual reading in mass miniature radiography, *American Review of Tuberculosis*, 61: 443, 1950
3. Yerushalmy, J. at all: An evaluation of the role of serial chest roentgenogram in estimating the progress of disease in patients with pulmonary tuberculosis, *American Review of Tuberculosis*, 64: 225, 1954.
4. Humerfelt, S.: An epidemiological study of high blood pressure, Bergen, University Press, 1963, p. 37
5. Thomas, A.J., Cotes, J. E. and Higgins, I.T.T.: Prevalence of coronary heart disease in elderly coal workers, *Lancet*, 1:414,1956
6. Thomas, A.J., Cochrane, A.L. and Higgins, I.T.T.: The measurement of the prevalence of ischemic heart disease, *Lancet*, 2: 540, 1958.