

# KLİNİK ARAŞTIRMALARDA İKİ ÖLÇÜM TEKNİĞİNİN UYUMUNU İNCELEMEDE KULLANILAN İSTATİSTİKSEL YÖNTEMLER

Yasemin Genç\* ❖ Durdu Sertkaya\*\* ❖ Selda Demirtaş\*\*\*

## ÖZET

Bazı klinik çalışmalarda, yeni bir ölçüm tekniği ile referans olarak kullanılan ölçüm tekniğinin uyum düzeyi çeşitli istatistiksel yöntemler kullanılarak incelenir. Genellikle ölçüm teknikleri arasındaki uyum düzeyini belirlemek için korelasyon ve regresyon analizi kullanılır. Fakat korelasyon ve regresyon analizi uyumu ölçmede doğru yöntemler değildir. Çalışmamızda bu yöntemlere alternatif olan ve grafiksel teknik ve basit hesaplamaları temel alan Bland-Altman yöntemi tanıtılmıştır.

**Anahtar Kelimeler:** Bland-Altman Yöntemi, Korelasyon Katsayısı, Metot Karşılaştırma Çalışmaları, Regresyon Analizi, Uyum Sınırları

## SUMMARY

### **Statistical Methods for Assessing Agreement Between Two Measurement Technique in Clinical Research**

In some clinical researches, degree of agreement between a new measurement technique and an established one is assessed by using a variety of statistical analysis. In general, correlation and regression analysis are commonly used for this purpose. However use of these analysis leads to misleading. In this study, an alternative approach, "Bland-Altman method", based on graphical techniques and simple calculation is described on a diagnostic measurement technique.

**Key Words:** Bland-Altman Method, Correlation Coefficient, Method Comparison Studies, Regression Analysis, Agreement Limits

Klinisyenler, bir parametreyi ölçmeye yarayan ve eskisine göre daha ucuz ve çabuk yanıt veren yeni bir yöntem bulunduğunda (örn. CX7 aleti kullanılarak yapılan direk HDL-C ölçümü) bu yöntemi, yaygın olarak kullanılan referans yöntemle (örn. MANUEL olarak yapılan çöktürücülü HDL-C ölçümü) karşılaştırarak yöntemlerin ne derece uyumlu sonuçlar verdiğini bulmak isterler. Burada uyum ile kastedilen iki yöntemden elde edilen ölçüm değerlerinin eşit olmasıdır. Aynı parametreyi ölçmek için kullanılan iki farklı yöntemin tüm denekler için tamamen aynı sonuçları vermesi pek olanaklı değildir. Fakat yeni yöntemin referans yöntemden ne kadar farklı sonuçlar verdiğini bul-

mak mümkündür. Eğer bu fark klinik yorumu etkileyebilecek düzeyde değil ise referans yöntem yerine yenisi ya da yöntemler değişimli olarak kullanılabilir(1).

Genellikle iki yöntem arasındaki uyumu bulmak için yapılan çalışmalarda korelasyon ve regresyon analizleri kullanılır. Uyumu araştırmada bu tekniklerin kullanılması aşağıdaki nedenlerden dolayı doğru değildir.

1. Korelasyonun derecesi örneklemedeki sonuçların dağılım genişliğine bağlıdır. Korelasyon, dağılım genişliği büyük olan örneklerde, dar olan örneklerle göre daha yüksek çıkar. Örneğin bu

\* Öğr. Gör. Dr., Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı

\*\* Öğr. Gör. Dr., Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü

\*\*\* Doç. Dr., Ankara Üniversitesi, Tıp Fakültesi, Acil Servis Biyokimya Laboratuvarı Sorumlusu

durum 30 ve 5.5 g/dl gibi dar dağılım genişliğine sahip Albimün ölçümleri için dezavantajken, 0 ve 43 gibi büyük dağılım genişliğine sahip ALT ölçümleri için avantajdır. Atkinson ve Nevill (1997) makalelerinde uyum düzeyleri tamamen eşit fakat dağılım genişlikleri farklı olan iki veri setinden elde edilen korelasyon katsayılarının çok farklı bulunduğunu örnek üzerinde göstermişlerdir.

2. Korelasyon katsayısının anlamlılığının testi, "iki yöntem arasında bağıntı yoktur" hipotezinin testidir. Aynı değeri ölçmek için dizayn edilen iki yöntemin bağıntılı (ilişkili) olup olmadığını test etmek gereksizdir.

3. İki yöntem zayıf uyuma sahip olduğu halde yüksek korelasyon gösterebilir. Örneğin A yöntemi, B yöntemi ile elde edilen değerlerin tam 2 katını veriyorsa, iki yöntemin sonuçları arasında tam bir korelasyon bulunurken zayıf bir uyumluluk elde edilecektir.

4. Doğrusal regresyon analizi yöntem karşılaştırma çalışmalarında kullanılan diğer bir tekniktir. Regresyon analizinin amacı, noktalara en yakın doğruyu çizebilmek için a ve b katsayılarını kestirmektir. a ve b katsayılarının anlamlılık testi O'a karşı yapılır. Fakat uyumu göstermede regresyon analizi kullanılacaksa regresyon eğrisi  $a=0$  ve  $b=1$ 'e karşı test edilmelidir. Yani iki yönteme ait değerler saçılım grafiğinde eşitlik çizgisi (line of equality) üzerinde olmalıdır. Paket programlarda a ve b'nin testinin O'a karşı yapıldıkları unutulmalıdır.

5. Regresyon analizi x bağımsız değişkeni ile y bağımlı değişkeni arasındaki ilişkiyi inceler. Uyumu göstermede regresyon analizi kullanılırken hangi yönteme ait verilerin x, hangilerinin y olarak kabul edilmesi gerektiğine dair bir bilgi yoktur.

6. Regresyon katsayısı da korelasyon katsayısı gibi dağılım genişliğinden etkilenir (1-5).

Bu problemler istatistik literatürlerinde birçok kez dile getirilmiştir. Bu analizlere alternatif bir yöntem Bland-Altman tarafından ilk olarak 1983 yılında verilmiştir.

## Yöntem

### Uyumun Ölçümü

Yöntem karşılaştırması çalışmasına iki yöntemden elde edilen ölçümlerin ortalamalarına karşı

farklarının saçılım grafiğinin çizilmesi ile başlanır. Farklara karşı ortalamaların grafiği ölçüm hataları ile gerçek değerler (gerçek değerler elde edilemediğinden onun en iyi kestiricisi olan ortalamalar) arasında olabilecek herhangi bir ilişkinin incelenmesine olanak sağlar. Ayrıca bu grafikten yan'ın (bias) (farkların sistematik olarak O'dan farklı olup olmadığının) ve hatanın (farkların ne kadar yaygın olduğunun) incelenmesi de mümkündür(4).

Farklar ve ortalamalar arasında ilişki yoksa iki yöntem arasındaki uyum farkların ortalaması ( $\bar{d}$ ) ve standart sapması (s) kullanılarak incelenebilir(1).

Eğer farklar normal dağılıma sahipse farkların sıfırın etrafında rasgele dağılması ve %95'inin " $\bar{d}-1,96s$  ile  $\bar{d}+1,96s$ " arasında olması beklenir. Bu durum altında ortalamalar ve farklar arasında ilişki olmadığı söylenebilir(1,4). Bu yöntemde,  $\bar{d}\pm 1,96s$  "uyum sınırları" olarak adlandırılır. Küçük örneklem için uyum sınırları bulunurken tablo değerini (n-1) serbestlik derecesi için t tablosundan yararlanarak bulmak daha doğru olacaktır. Uyum sınırları kısaca şu şekilde açıklanabilir. A yöntemi bir biyokimyasal değeri ölçmede rutin olarak kullanılan bir yöntem olsun. B ise A yöntemi yerine kullanılması düşünülen başka bir yöntem olsun. "B yöntemi kullanıldığında elde edilen ölçüm değeri, A yöntemi ile elde edilen değerden  $\bar{d}-1,96s$  kadar küçük,  $\bar{d}+1,96s$  kadar büyük bulunabilir". Bu sınırlar kabul edilebilir düzeyde midir? Bu sorunun istatistiksel bir cevabı yoktur. Buna araştırmayı planlayan kişi karar verebilir. Kabul edilebilir uyum sınırlarına veriler toplanmaya başlanmadan önce çalışma planlandığı zaman karar verilmelidir.

Eğer farklar ve ortalamalar arasında ilişki gözlenirse uyum sınırları yukarıda verilen yöntemle bulunamaz. Çünkü farkların gözlenen sınırları ortalama değerine bağlı olur. Bu durumda başvurulacak ilk yöntem logaritma dönüşümü uygulamaktır. Eğer bu yöntemle ilişki ortadan kaldırılabirirse uyum sınırları anti-logaritma kullanılarak tekrar hesaplanabilir. Aksi durumda, fark değerleri ancak grafikte gösterilebilir ve uyum sınırları istatistiksel bir değer gibi değil de kabaca tanımlayıcı bir değer gibi verilebilir(4).

### Tekrarlanabilirlik

Tekrarlanabilirlik, aynı denek üzerinde tekrar edilen ölçümler arasındaki değişimin bir ölçüsüdür. Bu ölçüt bir yöntemin kullanılabilir olup olmadığına karar vermede önemli bir role sahip olduğu gibi iki yöntemi karşılaştırırken de kullanılır. Tekrarlanabilirlik yöntemlerin uyumunu önemli ölçüde etkilemektedir. Eğer yöntemlerden biri zayıf tekrarlanabilirliğe sahipse iki yöntem arasındaki uyumun zayıf olması kaçınılmazdır. Eğer referans yöntemin tekrarlanabilirliği düşükse yeni yöntem çok iyi bir yöntem olsa da uyum yüksek çıkmayacaktır. Eğer iki yöntemin de tekrarlanabilirliği düşükse problem daha da büyüyecektir. Yöntem karşılaştırma çalışmalarında iki yöntem uyumlu çıkarsa bu yöntemlerden tekrarlanabilirliği yüksek olanın diğerine göre daha üstün olduğu söylenebilir.

Tekrarlanabilirlik katsayısı kurulan deney düzenine bağlı olarak değişik şekillerde hesaplanır. Örneğin sadece bir kişiden alınan kan örneği n defa çalışılırsa bu diziye ait varyasyon katsayısı bu tür bir deney düzeni için tekrarlanabilirlik katsayısı olarak kullanılır. İkinci bir yol n kişiden alınan kan örneklerinin her birini iki kez çalışmaktır. Bu tür bir deney düzeninde yaygın olarak kullanılan tekrarlanabilirlik katsayısı iki ölçüm arasındaki farklara ait standart sapmanın 2 katıdır(1).

Tekrarlanabilirlik çalışmalarında da yöntem karşılaştırma çalışmalarında olduğu gibi yaygın olarak korelasyon katsayısı kullanılmaktadır. Bland ve Altman daha önce belirtilen nedenlerden dolayı korelasyon katsayısı kullanılmasının yanlış olduğunu, bunun yerine yukarıda belirtilen yöntemlerin yanında kendi önerdikleri yöntemin de kullanılabileceğini belirtmişlerdir(1).

### Materyal

Bu çalışmada, Ankara Üniversitesi İbni-Sina hastanesine başvuran 80 hastanın serumlarında çöktürücü kullanılarak yapılan manuel HDL-C ile RA-XT ve CX7 aletleriyle direkt olarak ölçülen HDL-C sonuçları arasındaki uyum Bland ve Altman tarafından önerilen yöntemle incelenmiştir.

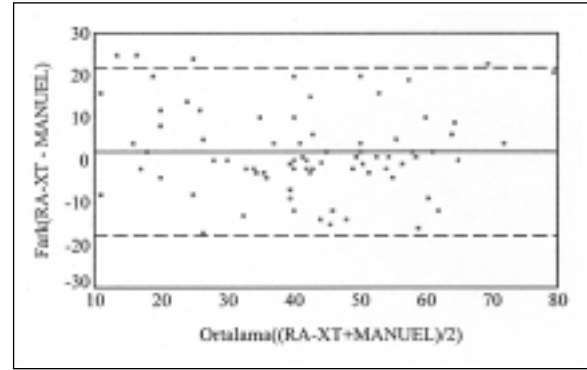
Rutin olarak kullanılan MANUEL HDL-C yönteminin zaman alıcı olması ve kişilerin diğer biyokimyasal değerlerinden etkilenmesi yeni ölçüm yöntemlerine gereksinim duyulmasına sebep ol-

muştur. Bu amaçla iki yeni ölçüm yöntemi geliştirilmiştir. Bunlardan biri RA-XT aleti ile diğeri de CX7 aleti ile direkt olarak yapılan HDL-C ölçüm yöntemidir. RA-XT ve CX7 aletleriyle elde edilen HDL-C ölçümlerinin çöktürücü kullanılarak MANUEL yolla elde edilen ölçümlere ne düzeyde uyum gösterdiği uyum sınırları kullanılarak incelenecektir. Uyum, klinisyenler tarafından kabul edilebilir düzeyde bulunur ise MANUEL yöntemin yerini daha çabuk sonuç veren yöntemlerden biri alacaktır. Ayrıca iki yeni yöntemin uyumu da uyum sınırları kullanılarak incelenecektir.

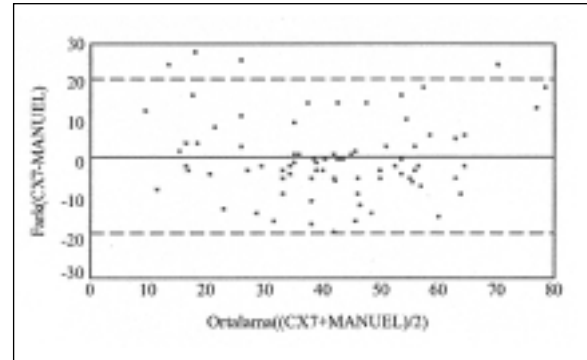
### Bulgular ve Tartışma

İncelemeye, iki yöntemin sonuçları arasındaki farkların ortalamalarına karşı çizilen saçılım grafiği ile başlanmıştır. Bu grafikler farkların sıfırın etrafında sistematik bir dağılım gösterip göstermediğinin ve hangi yaygınlıkta olduğunun incelenmesine olanak sağlar.

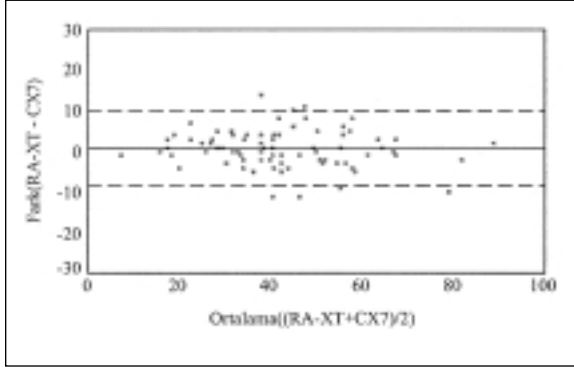
Aşağıda yöntem çiftlerine ait grafikler verilmiştir.



Şekil 1: RA-XT ve MANUEL yöntemlerine ait ölçümlerin ortalamaya karşı fark değerlerinin saçılım grafiği



Şekil 2: CX7 ve MANUEL yöntemlerine ait ölçümlerin ortalamaya karşı fark değerlerinin saçılım grafiği



**Şekil 3:** RA-XT ve CX7 yöntemlerine ait ölçümlerin ortalama karşı fark değerlerinin saçılım grafiği

Grafiklerden farkların sıfırın etrafında sistematik bir dağılım göstermediği ve farklarla ortalamalar arasında açık bir ilişki olmadığı görülmektedir. O halde yöntemler arası uyum, farkların ortalaması ve uyum sınırları bulunarak incelenebilir. Üç yöntemin ikiyeşerli olarak karşılaştırmalarına ait değerler Tablo 1’de verilmiştir.

Tablo 1’e göre analiz RA-XT aleti ile yapıldığında elde edilen HDL-C değeri, MANUEL olarak yapıldığında elde edilen değerden 18,4 birim düşük, 21,2 birim yüksek bulunabilir. Analiz CX7 aleti ile yapıldığında elde edilen HDL-C değeri, MANUEL olarak yapıldığında elde edilen değerden 17,7 birim düşük, 21,9 birim yüksek bulunur. İki direk yöntem karşılaştırılacak olursa, analiz RA-XT aleti ile yapıldığında elde edilen HDL-C değeri, CX7 aleti ile elde edilen değerden 8,3 birim düşük, 9,8 birim yüksek bulunabilir. Klinisyen bu sonuçları değerlendirerek yeni yöntemlerden birinin MANUEL yöntemin yerini alıp alamayacağına karar verebilir. Örneğimizde iki yeni yöntemin MANUEL yöntemle uyum düzeyleri

benzer bulunmuştur. Bu durumda CX7 ve RA-XT yöntemlerinden hangisi daha üstündür, hangisi kullanılmalıdır? Bu kararı vermede yöntemlerin tekrarlanabilirliği önem kazanır. Örneğimizde tekrarlanabilirlik varyasyon katsayıları bulunarak incelenmiştir. Buna göre varyasyon katsayıları HDL-C değerleri MANUEL olarak ölçüldüğünde 1.29, RA-XT aleti ile ölçüldüğünde 0.90, CX7 aleti ile ölçüldüğünde ise 0.79 olarak bulunmuştur. Örneğimizde tekrarlanabilirliği daha iyi olan CX7 yönteminin RA-XT yöntemine göre daha üstün olduğu söylenebilir.

### Sonuç

Yöntem karşılaştırma araştırmalarının korelasyon katsayısı ve regresyon analizi ile değerlendirilmesinin uygun olmadığı, çalışmamızda nedenleri ile açıklanmış, bunların yerine Bland ve Altman tarafından geliştirilen yöntemin kullanılmasının gerekliliği vurgulanmıştır. Bu yöntemin en önemli özelliği iki yöntemin ölçüm farklılıklarını objektif olarak ortaya koyması ve farklılıkların kabul edilebilirlik düzeyinin yorumunu klinisyenin görüşüne bırakmasıdır.

Neden bu amaçla kullanılması yanlış olduğu halde korelasyon katsayısı bu kadar yaygın olarak kullanılmaktadır? Bunun birinci nedeni veri yapısının bu analizleri kullanmaya uygun olması, ikincisi de alışkanlıktır. İki yöntemle ait değerlerin saçılım grafiği çizildiğinde bu grafik korelasyon katsayısı kullanıldığı durumlardaki grafiği hatırlatır. Ayrıca makalelerde korelasyon ve regresyon analizlerinin kullanılmış olması araştırmacıların bu yöntemleri rahatlıkla kullanmalarına sebep olmuş ve bu yanlış kullanımlar yaygınlaşmıştır.

**Tablo 1:** Üç yöntem çiftine ait  $\bar{d}$  ve uyum sınırları

Metodlar	n	$\bar{d}$	S. Sapma	Uyum sınırları	
				$\bar{d} \pm 1,96*S.Sapma$	
MANUEL - RA-XT	80	1,38	10,09	-18,4	21,2
MANUEL - CX7	80	2,14	10,13	-17,7	21,9
CX7 - RA-XT	80	0,76	4,61	-8,3	9,8

**KAYNAKLAR**

1. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement Lancet 1986; I: 307-10
2. Atkinson G, Nevill A. Comment on the use of concordance correlation to assess the agreement between two variables. Biometrics 1997; 775-777
3. Altman DG, Bland JM. Measurement in medicine: The analysis of comparison studies. Statistician 1983; 32:307-317
4. Hollis S. Analysis of method comparison studies JIFCC 1997; 9
5. Pollack MA, Jefferson SG, Kane JW, et al. Method comparison - a different approach. Ann Clin Biochem 1992; 29:556-560

