

İki Ölçüm Tekniğinin Uyumunu İncelemede Kullanılan Grafıksel ve Regresyon Yöntemleri: Uygulamalı ve Senaryolara Dayalı Bir Çalışma

Graphical and Regression Methods Which are Used to Examine Agreement of Two Measurement Techniques: An Applied and Scenarios Based Study

Selen YILMAZ IŞIKHAN,^a
Mustafa KILIÇKAP,^b
C. Reha ALPAR^c

^aHacettepe Üniversitesi,
Sosyal Bilimler Meslek Yüksekokulu,
^bKardiyoloji AD,
Ankara Üniversitesi Tıp Fakültesi,
^cBiyostatistik AD,
Hacettepe Üniversitesi Tıp Fakültesi,
Ankara

Geliş Tarihi/Received: 17.08.2012
Kabul Tarihi/Accepted: 10.12.2012

Bu çalışma XIII. Ulusal Biyoistatistik Kongresi (12-14 Eylül 2011, Kızılcahamam, Ankara)'nde sunulmuştur.

Yazışma Adresi/Correspondence:

Selen YILMAZ IŞIKHAN
Hacettepe Üniversitesi,
Sosyal Bilimler Meslek Yüksekokulu,
Ankara,
TÜRKİYE/TURKEY
seleny@hacettepe.edu.tr

ÖZET Amaç: Çalışmanın amacı, özellikle ölçüm teknikleri arasında oransal yanlışlık (bias) bulunduğu, uyum sınırlarına alternatif olan regresyona dayalı güven sınırlarını elde etmek, farklı uyum analiz yöntemlerinin değişik senaryolu veri setleri üzerinde kullanımlarını özetlemektir. **Gereç ve Yöntemler:** Çalışmada koroner anjiyografi yapılan 88 hastaya ait 200 damar darlığı, biri uzman, diğeri 2 aylık deneyimi olan bir araştırma görevlisi tarafından görsel olarak "% darlık" şeklinde değerlendirilmiş ve bu değerlendirmeler arasındaki uyum incelenmiştir. Orijinal veri, farklar ve ölçüm değerleri arasında homojen saçılmış oransal yanlışlık, heterojen saçılmış oransal yanlışlık ve eğrisel ilişki durumlarını içerecek şekilde modifiye edilmiştir. Daha sonra bu dört farklı senaryo üzerinde farkların analizi, regresyona dayalı güven sınırları, model I regresyon yöntemleri (En Küçük Kareler (EKK) ve Ağırlıklı En Küçük Kareler (AEKK)) ve model II regresyon yöntemlerinin (Deming, Ağırlıklı Deming ve Passing Bablok) performansı karşılaştırılmıştır. **Bulgular:** Farklar metodu bir ölçüm tekniğinin diğeri yerine güvenilir bir şekilde kullanılabilir olup olmadığı hakkında klinik yoruma dayalı bilgi verirken, model II regresyon yöntemleri karşılaştırılan ölçüm teknikleri arasındaki sistematik yanlışlığı ortaya çıkarmada başarılıdır. EKK ve AEKK yöntemleri ise, ölçüm tekniklerinin her birini ayrı ayrı bağımlı değişken olarak iki farklı denklem tahmini vermekte, bu denklemler ise α ya da β güven sınırları (sabit ya da oransal yanlışlık) açısından birbirleriyle çelişmektedir. **Sonuç:** Farklar ve ortalamalar arasında oransal yanlışlık bulunmuyorsa klasik % 95 uyum ya da tolerans sınırları değerlendirilebilir. Farklar ve ortalamalar arasındaki ilişki değişen varyanslı ise, ayrıca logaritmik dönüşüm ile oransal yanlışlık giderilemez ise ortalamalara göre farkların regresyon doğrusu/eğrisi etrafında regresyona dayalı güven sınırları elde edilmektedir. Ölçüm teknikleri arasında oransal yanlışlık bulunması durumunda diğer bir seçenek, hem sabit hem de oransal yanlışlığı belirleyebilen model II regresyon yöntemlerini uygulamaktır.

Anahtar Kelimeler: Bland altman; regresyona dayalı güven sınırları; oransal yanlışlık; deming; Passing Bablok yöntemi

ABSTRACT Objective: The aim of this study is to obtain the limits of confidence alternative to limits of agreement (regression based) especially when there is proportional bias between measurement techniques and to summarize the application of different agreement analysis methods on data sets with various scenarios. **Material and Methods:** In this study, coronary angiographically detected 200 stenosed segments of 88 patients were visually graded as percent stenosis by an experienced staff and a fellow with 2-month experience in coronary angiography laboratory, then were examined in terms of agreement. The original data was modified so that proportional bias with homogenous scattering, proportional bias with heterogeneous scattering and curvilinear relationship cases between differences and measured values are included. Then, on the four different scenarios; the performances of analysis of differences, regression-based confidence limits, model I regression methods (Ordinary Least Squares (OLS), Weighted Least Squares (WLS) and model II regression methods (Deming, Weighted Deming and Passing Bablok) were compared. **Results:** While method of differences provides information based on the clinical interpretation about whether a measurement technique can be used safely instead of the other, model II regression methods are successful in revealing the systematic bias between the compared measurement techniques. OLS and WLS methods, taking each one of the measurement techniques dependent variable separately, give the two different estimated equations but these equations in terms of α or β in the confidence limits (fixed or proportional bias) conflict. **Conclusion:** If there is no proportional bias between differences and the mean values, classical 95% limits of agreement or 95% limits of tolerance can be evaluated. If the relationship between differences and mean values is in the form of heteroscedasticity, if also the proportional bias couldn't be eliminated with logarithmic transformation then limits of confidence based on regression around regression line/curve of differences according to mean values should be obtained. Another choice in the presence of proportional bias between measurement techniques is applying model II regression methods revealing both fixed and proportional bias.

Key Words: Bland Altman; limits of confidence based on regression; proportional bias; deming; Passing Bablok method

Tıpta tanı amaçlı kullanılan bazı tetkiklerde, özellikle de görüntüleme ile ilgili tanı yöntemlerinde çap, hacim gibi ölçümler yapılmaktadır. Bu ölçümler bazen en güvenilir olan “altın standart” tanı yöntemiyle, bazen de altın standart olmasa bile güvenilir olduğu kabul edilen daha ucuz veya daha pratik yöntemlerle yapılmaktadır. Yeni bir tanı yönteminin kullanıma girebilmesi için geçerli ve güvenilir olduğunun gösterilmesi gereklidir. Bu amaçla da yeni yöntem ile elde edilen ölçümlerin eski yöntemlerle elde edilen ölçümlerle uyumu değerlendirilir. Yeni ölçüm tekniği ile standart ölçüm tekniği arasındaki farklılık klinik yorumu değiştirecek boyutta değil ise, yeni yöntem eskisi yerine, ya da iki yöntem değişimli olarak kullanılabilir.^{1,2}

İki yöntem ile yapılan ölçümler arasındaki farklılık rastgele bir hatadan kaynaklanabileceği gibi sistematik yanlılıktan da kaynaklanabilir. Sürekli biyolojik bir değişkeni ölçen iki ölçüm tekniğinin karşılaştırılmasındaki amaç, benzerliklerden ziyade sistematik yanlılığı ortaya çıkarmaktır. Sistematik yanlılık, sabit yanlılık ve oransal yanlılık olmak üzere iki şekilde olabilir. Sabit yanlılığa, bir ölçüm tekniğinin tüm ölçüm aralığında diğer ölçüm tekniğinden sabit bir miktarda daha düşük ya da daha yüksek değerler vermesi şeklinde bir örnek verilebilir. Oransal yanlılığa ise bir ölçüm tekniğinin ölçülen değişkenin seviyesi ile orantılı olarak diğer ölçüm tekniğinden daha yüksek ya da daha düşük değerler vermesi olarak bir örnek verilebilir.³

Literatürdeki çoğu çalışmada iki ölçüm tekniği arasındaki uyumun bir ölçüsü olarak korelasyon katsayısı, saçılım grafikleri ve En Küçük Kareler (EKK) regresyon yöntemleri kullanılmaktadır. Yüksek korelasyon, iki ölçüm tekniğinin mükemmel uyuma sahip olduğunu göstermez. Noktalar herhangi bir regresyon doğrusu üzerinde yer aldığı müddetçe mükemmel bir korelasyon elde edilirken, 45°’den geçen eşitlik çizgisi üzerinde yer alırsa mükemmel bir uyuma sahip olunur.⁴

Klasik saçılım grafikleri ile de uyumu değerlendirmek yanıltıcı olabilmektedir. Bu nedenle Bland ve Altman tarafından önerilen ölçümler arasındaki farkların ortalamalara karşı saçılım grafikleri ölçüm tekniği farklılıklarını değerlendirmede

daha sık olarak kullanılan bir yöntem olmaya başlamıştır.⁵

Doğrusal regresyon analizinde (model I regresyon) y değerlerinin regresyon doğrusundan dikey sapmalarının kareler toplamı minimize edilir. Ancak uyum çalışmalarında hem x hem de y değerleri belirli bir ölçüm hatası içerdiğinden, bu yöntemin kullanımı uygun değildir. Dolayısıyla bir metodu diğerine göre ayarlamak ya da sistematik ölçüm hatalarını ortaya çıkarmak amacıyla model II regresyon yöntemleri tercih edilmelidir.^{3,6}

Bu çalışmada Bland-Altman grafikleri ve model II regresyon yöntemlerinin ölçümler arasındaki uyumu değerlendirmede kullanımları, değişik şekillerde hata tipleri olan veri setleri kullanılarak özetlenmeye çalışılmıştır.

GEREÇ VE YÖNTEMLER

FARKLARIN ANALİZİ VE KLASİK UYUM SINIRLARI

İki ölçüm tekniği ile elde edilen ölçümlerin analizine, ortalamalara karşı farkların saçılım grafiği çizilerek başlanır. Doğrusal regresyondaki artıkların analizine benzer şekilde, farklar ve ortalamalar arasında ilişki olmaması istenir. Bu durumda, farklar grafiğinde uyum kaybı, farkların ortalaması \bar{d} ve standart sapması (S_d) kullanılarak özetlenir.¹

Farklar normal dağılım gösteriyorsa, farkların %95’i $\bar{d} - 1,96S_d$ ile $\bar{d} + 1,96S_d$ arasında yer almaktadır. Altman ve Bland, tüm örneklem genişlikleri için daha uygun uyum sınırları olarak $\bar{d} \pm (t_{0,05;n-1}) (S_d)$ ’nin kullanılmasını önermişlerdir.

Uyum sınırları, bir popülasyondan çekilen tek bir örneklem için elde edilen tahmin değerleridir. İkinci bir örneklem farklı uyum sınırları verecektir. Farklar yaklaşık olarak normal dağılıyorsa uyum sınırları tahminlerinin kesinliği, alt ve üst uyum sınırlarının %95 güven (tolerans) sınırları ile incelenir.⁶

Uyum sınırları $\bar{d} \pm 1,96S_d$ ’nin standart hataları (SH) yaklaşık olarak $1,71 \frac{S_d}{\sqrt{n}} = 1,71 SH (\bar{d})$ bulunur. Uyum sınırları için %95 güven aralıkları, n-1 serbestlik dereceli t dağılımının uygun noktası bulunarak (gözlenen değer $\pm t \times SH$) biçiminde hesaplanabilir.¹

İki ölçüm tekniği arasında sabit bir yanlılık yoksa, bu durumda farkların (X-Y) ortalaması sıfır olmalıdır. Bu hipotez ise tek-örneklem t testi ile değerlendirilir. Diğer bir seçenek ise fark ortalaması için %95 güven sınırlarının sıfır içerip içermediğine bakmaktır.

Oransal bir fark yoksa *farkların ortalamalar üzerindeki regresyonu sıfır eğimine sahip olmalıdır*. Bunun yanı sıra, Pearson'ın korelasyon katsayısı için $\rho=0$ hipotezi test edilir.

REGRESYONA DAYALI GÜVEN SINIRLARI

Yukarıda farkların analizi bölümünde anlatılan %95 uyum ve tolerans sınırlarının önemli bir varsayımı, farkların ve farkların standart sapmasının ölçüm düzeyinden (iki ölçümün ortalamasına ait değerlerden) bağımsız ve sabit olduğudur. Ancak bu varsayımlar her zaman sağlanamayabilir. Bu varsayımların sağlanıp sağlanmadığı, *farkların ortalamalar üzerindeki regresyon eğiminin (β) ya da Pearson korelasyon katsayısının (ρ), sıfıra karşı testi ile belirlenir. Regresyon doğrusunun eğiminin sıfırdan farklı olması halinde farkların büyüklüğünün ortalamaların düzeyine bağlı olduğu, yani oransal yanlılığın bulunduğu sonucuna varılır.⁷ Oransal yanlılığın varlığında uyum sınırlarını belirlemek için önerilen iki farklı çözüm:*

(1) Veri dönüşümü

(2) Doğrusal regresyon doğrusu etrafında *simetrik güven sınırlarının* oluşturulması.

Logaritma dönüşümü, farklar ve ölçüm büyüklüğü arasındaki ilişkiyi ortadan kaldıramadığı durumda uyum sınırları analizi, *zayıf ölçüm tekniklerinin kabulüne* yol açabilir.

Bu gibi durumlarda güven sınırları *en iyi doğrunun her iki tarafında simetrik düzenlenen hiperbol* biçiminde elde edilebilir.⁶ Bu sayede logaritmik dönüşümlü sınırlara geri dönüşüm uygulamak yerine orijinal verilere dayalı sınırlar elde edilebilmektedir.⁸

Doğrusal regresyon yönteminde aşağıdaki adımlar izlenerek farkların güven aralığı belirlenebilir.^{6,9}

x ve y, iki ölçüm tekniğinden n gözlem için elde edilen değerler iken ortalama ve fark vektörleri aşağıdaki gibi bulunur:

$$\text{ortalama} = \frac{x + y}{2}, \text{ fark} = x - y$$

1- Farkların ortalamalar üzerinde doğrusal regresyonu oluşturulur:

$$\text{Fark} = a + b \times \text{ortalama} \quad (1)$$

2- Gözlenen ve beklenen değerler arasındaki farklar biçiminde artıklar elde edilir.

3- Artıkların mutlak değerleri alınır.

4- Mutlak artıkların ortalamalar üzerindeki EKK regresyonu oluşturulur. Yani:

$$\text{Mutlak Artık} = c + d \times \text{ortalama} \quad (2)$$

5- "d" Eğim katsayısı anlamlı ise, eşitlik (2)'deki regresyon katsayıları $\sqrt{\pi/2}$ 1.2533 ile çarpılarak farkların standart sapmasının kestirimi için (3) denkleminde elde edilir (Buradaki $\sqrt{\pi/2}$ çarpanı, normal dağılım gösteren farkların mutlak değerlerinin ortalaması olan $\sigma\sqrt{2/\pi}$ 'den gelmektedir). Buna göre:

$$S_d = 1.2533 \times c + (1.2533 \times d) \times \text{Ortalama} \quad (3)$$

6- Böylece, fark ortalaması ve standart sapması (1) ve (3) eşitlikleri yardımıyla tahmin edilerek %95 güven sınırları elde edilir.

$$\text{Üst ve alt sınırlar: } \text{eşitlik}(1) \pm 1.96 \times \text{eşitlik}(3) \quad (4)$$

Bu güven ya da kestirim sınırları, aynı popülasyondan çekilecek yeni örneklemelere ait ölçüm değerlerinin %95'inin içerisinde bulunacağı aralığı belirtmektedir.

MODEL II REGRESYON YÖNTEMLERİ

Bazen, özellikle bir ölçüm tekniğinin başka bir ölçüm tekniğine göre ayarlanmasında, çok uygun olmamakla birlikte EKK esasına dayalı model I regresyon analizi yapılmaktadır. Bu amaçla model I regresyon analizinin uygun olmamasının nedenlerinden biri, iki ölçüm tekniğinin karşılaştırılmasında genellikle bağımlı ve bağımsız değişken ayırımının yapılamaması, bu nedenle de hangisinin bağımlı değişken olacağına göre değişen iki ayrı regresyon doğrusunun elde edilmesidir.

Diğer bir neden ise EKK yönteminin bir varsayımının ihlal edilmesidir. EKK yönteminde x ölçümünün sabit her bir değeri için y ölçümündeki değişim dikkate alınır. Yani değişim (hata), ölçümlerin sadece birinde olur. Halbuki ölçüm tekniği karşılaştırmalarında her iki ölçüm tekniğinde de rastgele hatalar oluşması söz konusudur. Bu nedenlerden dolayı iki ölçüm tekniğinin karşılaştırılmasında model II regresyon yöntemleri daha uygun görülmektedir.^{3,10}

İki ölçüm tekniği arasındaki doğrusal regresyon denkleminde oransal sistematik yanlışlık “β=1”, sabit sistematik yanlışlık; “α=0” hipotezleri test edilerek belirlenebilir.

EKK yönteminde regresyon doğrusuna düşey yönde olan uzaklıklar en aza indirilmeye çalışılırken, model II regresyon yöntemlerinde regresyon doğrusuna hem yatay hem de dikey yöndeki uzaklıkların çeşitli fonksiyonları en küçük yapılır. Böylelikle her iki ölçüm tekniğindeki rastgele hatalar dikkate alınmış olur.^{3,10}

Özellikle x ve y ölçüm tekniklerinden hangisinin bağımlı değişken olarak alınacağını belirsiz olduğu ve her iki ölçüm tekniğinin rastgele hata içerdiği durumlar için tasarlanan model II regresyon yöntemlerinin yaygın olarak kullanılan iki türü vardır:

- *En Küçük Çarpımlar (Deming) Regresyonu*
 - *Ağırlıklı Deming Regresyon*
- *Passing - Bablok Regresyon*

Deming regresyon analizi, regresyon doğrusundan hem dikey yöndeki sapmaların hem de yatay yöndeki sapmaların normal dağılıma uyduğunu varsayar.¹¹ Deming regresyon yönteminde minimize edilmek istenilen hata kareler toplamı (HKT) aşağıdaki gibidir:

$$HKT = \sum\{(x_i - X_i)^2 + \lambda(y_i - Y_i)^2\} \quad (5)$$

$$\lambda = \frac{S_{ex}^2}{S_{ey}^2} \quad (6)$$

x ve y ölçümlerinin varyanslarının oranı olan λ değeri, sapmaların kareler toplamlarını doğru üzerinde minimize edecek açının belirlenmesini

sağlar. Yani x ve y ölçümlerine ait hataların kareler toplamını eş zamanlı olarak en küçük yapar. Bu metod, En Küçük Çarpımlar (EKÇ) regresyonu terimi ile de anılır, çünkü EKÇ regresyonunda x ve y ölçümlerinin doğrudan yatay ve dikey uzaklıklarının çarpımlar toplamı minimize edilir.^{3,11}

Sabit sistematik yanlışlık durumunda, yani sistematik farklar ölçüm düzeyinden bağımsız iken Deming regresyon uygundur.^{3,6} Ancak oransal yanlışlığın varlığında Deming regresyonun ağırlıklı biçimi gerçekleştirilir. Ağırlıklı Deming regresyon yönteminde, doğruya belli bir açıdaki ağırlıklı sapmaların kareler toplamı en küçük yapılır.

$b_{y,x}$ ve $b_{x,y}$, sırasıyla y'nin x üzerindeki ve x'in y üzerindeki Ağırlıklı En Küçük Kareler (AEKK) regresyon katsayıları iken regresyon katsayıları (7) ve (8) eşitlikleri yardımıyla elde edilir.^{3,11}

$$b' = \sqrt{(b_{y,x})(1/b_{x,y})} \quad (7)$$

$$a' = \bar{y} - b'\bar{x} \quad (8)$$

Passing ve Bablok, verinin dağılımı hakkında hiçbir özel varsayımın bulunmadığı, hem x hem de y ölçüm tekniğindeki hataları dikkate alan *sağlam* bir regresyon yöntemi önermişlerdir. Regresyon doğrusunun eğimi, tüm olası eğimlerin ortancası olarak hesaplanır. Bu yöntemin de sonuçları, referans ölçüm tekniğinin x (bağımsız değişken) ve test ölçüm tekniğinin y (bağımlı değişken) olarak atanmasından bağımsızdır.¹²

Hem Deming yöntemi hem de Passing-Bablok yönteminde, α için güven aralığı sıfırı içermezse karşılaştırılan yöntemlerin en az bir sabit değer kadar farklı olduğu söylenir. β için güven aralığı 1'i içermediği durumlar yöntemler en az oransal (oransal sistematik yanlışlık) bir değer kadar farklı olur.^{12,13}

GEREÇ VE YÖNTEMLER

Bu çalışmada, Ankara Üniversitesi Tıp Fakültesi Kardiyoloji Anabilim Dalında koroner anjiyografi yapılan ardışık 88 hastadaki 200 damar darlığına ilişkin değerler kullanıldı. Koroner anjiyografide koroner arterlerdeki darlık derecesi pratikte çoğu kez görsel olarak değerlendirilmekte ve darlık derecesi yüzde olarak ifade edilmektedir. Veri setini

TABLO 1: İki değerlendiricinin orijinal verilerine ait farkların analiz sonuçları.

Orijinal Veri (X - Y) ile (X + Y) / 2 için kurulan Hipotezler	Test istatistiği	Değer	Serbestlik Derecesi	P Değeri
$H_0 : \bar{D} = 0$ $H_1 : \bar{D} \neq 0$	t	5,656	199	<0,001
\bar{D} için % 95 güven aralığı		2,217 – 4,592		

oluşturan çalışma koroner anjiyografi laboratuvarında 2 ay çalışmış bir araştırma görevlisi ile deneyimli bir uzmanın görsel olarak değerlendirdikleri darlık dereceleri arasında uyumun olup olmadığını belirlemek üzere yapıldı.

Çalışmada, “İnsan” ögesinin içinde bulunduğu tüm çalışmalar Helsinki Deklerasyonu 2008 prensiplerine (<http://www.wma.net/en/30publications/10policies/b3/index.html>) uygun olarak yapıldı.

Oluşturulan veri seti üzerinde gözlemciler arası uyum farklı yöntemlerle değerlendirilmiştir. Bazı hastalara ait kümelenmiş veriler, uygulama amaçlı, kümeleme etkisi dikkate alınmaksızın analiz edildi.

Orijinal veriler daha sonra, farklar ile ortalamalar arasındaki üç farklı ilişki durumunu (hiperbol, megafon ve eğrisel) temsil edecek şekilde modifiye edilmiştir. Regresyona dayalı güven sınırları sırasıyla hiperbol (Oransal fakat sabit varyanslı), megafon (V biçimli, oransal fakat artan varyanslı) ve eğrisel biçimdeki farkların grafikleri için elde edilmiştir. Ölçümler, farkların analizi, EKK, AEKK ve model II regresyon yöntemleri olan Deming, Ağırlıklı Deming ve Passing-Bablok regresyon yöntemleri yardımıyla da ayrıca değerlendirilmiştir.

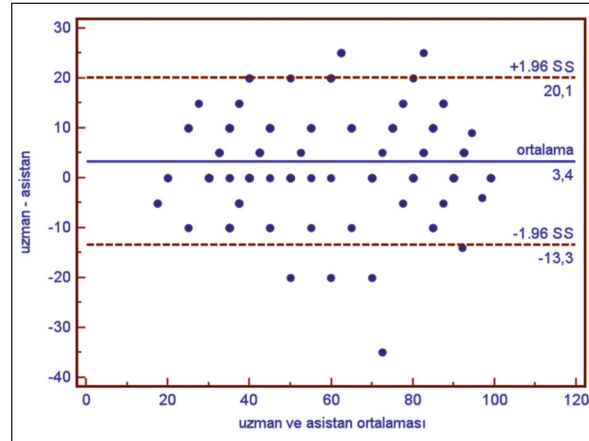
İstatistiksel hesaplamalar ve grafikler, Medcalc 9.3.9.0, SPSS 15.0 ve Excel eklentisi Analyse-it yazılımlarında gerçekleştirilmiştir.

BULGULAR

Senaryo 1: Orijinal Verilerin Fark ve Ortalamaları Arasında Homojen Dağılımın Mevcut Olduğu Durum

İncelemeye, iki değerlendiricinin orijinal sonuçları arasındaki farklar grafiği ve farkların analizi ile başlanılmıştır (Şekil 1, Tablo 1).

Şekil 1’den, farkların sıfır etrafında homojen bir dağılım gösterdiği ve farklarla ortalamalar ara-



ŞEKİL 1: Uzman ve 2 aylık deneyimli değerlendiricilere ait ölçümlerin ortalamaya karşı farklar grafiği.

sında açık bir ilişki olmadığı görülmektedir. Farklar ile ortalamalar arasındaki korelasyon katsayısı önemsizdir ($p=0.330$). Bu bulgu, Şekil 1’i destekler niteliktedir.

Farkların analiz sonuçları ise Tablo 1’de verilmiştir. Şekil 1 ve Tablo 1’den stajyer gözlemcinin uzmana göre darlık derecelerini ortalama %3,4 kadar daha düşük belirlediği görülmektedir ($t=5.656$, $p<0.001$). Bu bulgular sistematik olarak sabit bir yanlılığın bulunduğu, ancak oransal bir yanlılığın bulunmadığı anlamına gelmektedir.

Tablo 2 ve Tablo 3’de sırasıyla model I ve model II regresyon sonuçları verilmiştir. EKK sonuçları, hem sabit yanlılık hem de oransal yanlılık açısından tutarsız güven sınırları vermiştir. Deming ve Passing-Bablok yöntemlerine göre ölçüm teknikleri arasında hem sabit, hem de oransal yanlılık bulunmamıştır.

Senaryo 2: Farklar ve Ortalamalar Arasında Oransal ve “Sabit Varyanslı” İlişkinin Mevcut Olduğu Durum

Orijinal veriler saçılım grafiklerinden yararlanarak farklar ve ortalamalar arasında hiperbol bir ilişki elde edilecek şekilde değiştirilmiştir.

TABLO 2: İki değerlendiricinin orijinal verilerine ait EKK regresyonu sonuçları.

Yöntem	a	%95 GA	b	%95 GA
(1) $E(y)=a+b(X)$	(2) $E(X)=a+b(y)$			
EKK (1) denklemi	5,76	2,74-8,78	0,96	0,90-1,01
EKK (2) denklemi	1,76	-1,28-4,79	0,91	0,86-0,96

TABLO 3: İki değerlendiricinin orijinal verilerine ait model II regresyon yöntemi sonuçları.

Yöntem	a	%95 GA	b	%95 GA
Deming	2,04	-0,55/4,65	1,02	0,97-1,07
Passing - Bablok	0,00	0,00-0,00	1,00	1,00-1,00

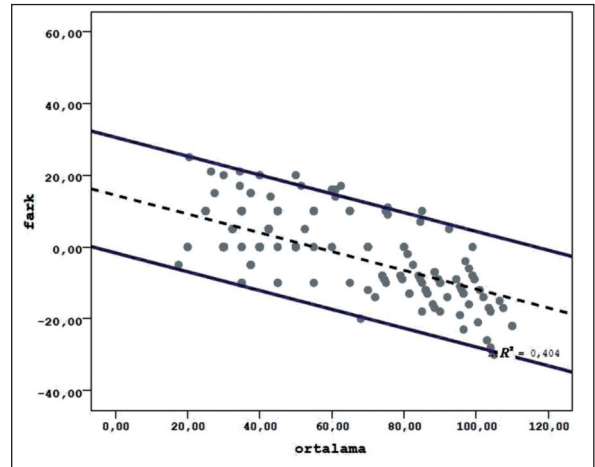
TABLO 4: İki değerlendiricinin verilerine ait farkların analiz sonuçları.

Orijinal Veri (X - Y) ile (X + Y) / 2 için kurulan Hipotezler	Test istatistiği	Değer	Serbestlik Derecesi	P Değeri
$H_0: \rho = 0$ $H_1: \rho \neq 0$	r	-0,636	200	<0,001
$H_0: \bar{D} = 0$ $H_1: \bar{D} \neq 0$	t	-0,746	199	0,456
\bar{D} için %95 güven aralığı		-2,03 – 0,91		
Log Dönüşümlü Veri				
log(X-Y) ile log (X + Y) / 2 için kurulan Hipotezler				
$H_0: \rho' = 0$ $H_1: \rho' \neq 0$	r	-0,513	200	<0,001
$H_0: \bar{D}' = 0$ $H_1: \bar{D}' \neq 0$	t	2,586	199	0,010
X/Y için % 95 güven aralığı		1,009-1,074		

Aşağıda modifiye edilmiş veriler üzerinde yöntem çiftleri için hesaplanan güven sınırı grafikleri, farkların analizi, model I ve model II regresyon yöntemi sonuçları verilmiştir.

Ortalama ve farklar arasındaki hiperbol biçimindeki ilk senaryo için farkların analizi gerçekleştirilmiş ve Tablo 4'te verilmiştir. Sonuçlara göre ölçüm teknikleri arasında oransal yanlılık bulunmakta ($r = -0,636$, $p < 0,001$), sabit yanlılık bulunmamaktadır.

Her iki ölçüme logaritma dönüşümü uygulanarak farklar analiz edildiğinde oransal yanlılık giderilememiş ($r = -0,513$, $p < 0,001$), bunun yanı sıra sabit yanlılığın varlığı da tespit edilmiştir ($t = 2,586$, $p = 0,010$). Oransal yanlılık giderilemediği durumda elde edilecek regresyona dayalı güven sınırları Şekil 2'de verilmiştir.

**ŞEKİL 2:** Farklar ve ortalamalar arası regresyona dayalı hiperbolik güven sınırları.

Hiperbol biçiminde modifiye edilmiş veri için farkların analizi (Tablo 4) sonucunda, sabit yanlılık tespit edilememiş ancak model II regresyon yön-

TABLO 5: İki değerlendiricinin verilerine ait model I regresyon yöntemi sonuçları.

Yöntem	a	%95 GA	b	%95 GA
Denklemler: (1) $E(y)=a+b(x)$ (2) $E(x)=a+b(y)$				
EKK (1) denklemi	14,72	12,52-16,93	0,73	0,70-0,77
EKK (2) denklemi	-12,35	-15,85-8,85	1,23	1,17-1,28
AEKK (1) denklemi	18,76	16,59-20,92	0,64	0,59-0,70
AEKK (2) denklemi	-4,51	-7,43-1,59	1,08	1,01-1,14

TABLO 6: İki değerlendiricinin verilerine ait model II regresyon yöntemi sonuçları.

Yöntem	a'	%95 GA	b'	%95 GA
Deming	-16.10	-19,63/-12,56	1,29	1,23-1,35
Ağırlıklı Deming	-16.47	-21,47/-11,77	1,30	1,21-1,38
Passing-Bablok	-12.28	-16,25/-8,00	1,25	1,20-1,32

TABLO 7: İki değerlendiricinin orijinal ve logaritmik verilerine ait farkların analiz sonuçları.

Orijinal Veri (X -Y) ile (X + Y) / 2 için kurulan Hipotezler	Test istatistiği	Değer	Serbestlik Derecesi	P Değeri
$H_0: \rho = 0$ $H_1: \rho \neq 0$	r	0,442	200	$p < 0,001$
$H_0: \bar{D} = 0$ $H_1: \bar{D} \neq 0$	t	8,54	199	$p < 0,001$
\bar{D} için % 95 güven aralığı		6,84-10,95		
Log Dönüşümlü Veri log(X-Y) ile log (X + Y) / 2 için kurulan Hipotezler				
$H_0: \rho' = 0$ $H_1: \rho' \neq 0$	r	0.314	200	$p < 0,001$
$H_0: \bar{D}' = 0$ $H_1: \bar{D}' \neq 0$	t	8.517	199	$p < 0,001$
X/Y için % 95 güven aralığı		1.110-1.183		

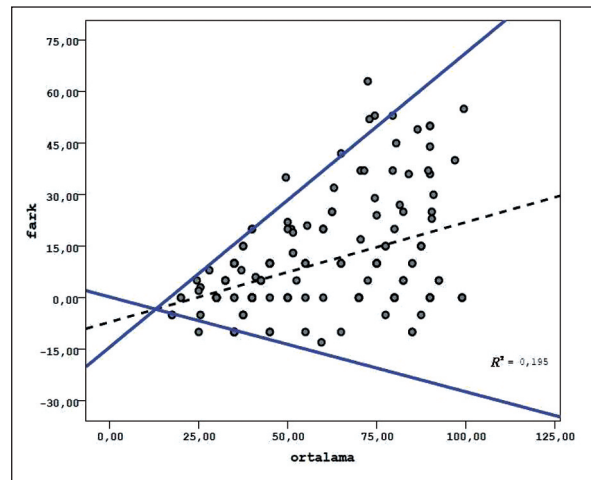
temleri (Tablo 6) hem sabit hem de oransal yanlılığı belirleyebilmiştir.

Ancak Tablo 5 sonuçlarında EKK ve AEKK yöntemleri sabit yanlılık (a katsayısı için güven aralıkları) açısından tutarsız sonuçlar vermiştir.

Model I regresyon sonuçları (EKK ve AEKK) ile elde edilen regresyon eğimleri (Tablo 5), model II regresyon yöntemleriyle elde edilen regresyon eğimlerinden daha küçük bulunmuştur (Tablo 6).

Senaryo 3: Farklar ve Ortalamalar Arasında Oransal ve "Artan Varyanslı" İlişkinin Mevcut Olduğu Durum

Orijinal veriler saçılım grafiklerinden yararlanarak farklar ve ortalamalar arasında megafon (V

**ŞEKİL 3:** Farklar ve ortalamalar arası regresyona dayalı megafon (V-biçimli) güven sınırları.

TABLO 8: İki değerlendiricinin verilerine ait model I regresyon yöntemi sonuçları.

Yöntem	a	%95 GA	b	%95 GA
Denklemler: (1) $E(y)=a+b(x)$ (2) $E(x)=a+b(y)$				
EKK (1) denklemi	4,28	-1,23-9,78	1,09	0,99-1,19
EKK (2) denklemi	12,65	8,80-16,49	0,64	0,58-0,70
AEKK (1) denklemi	-2,46	-6,95-2,02	1,23	1,12-1,35
AEKK (2) denklemi	10,10	7,88-12,32	0,68	0,63-0,74

TABLO 9: İki değerlendiricinin verilerine ait model II regresyon yöntemi sonuçları.

Yöntem	a'	%95 GA	b'	%95 GA
Deming	5,11	1,30-8,93	0,76	0,68-0,84
Ağırlıklı Deming	6,33	3,12-9,36	0,74	0,69-0,79
Passing- Bablok	6,00	0,00015-8,82	0,80	0,70-0,88

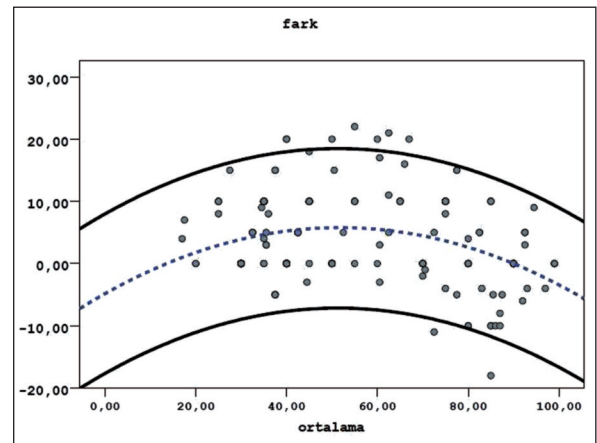
biçiminde) bir ilişki elde edilecek şekilde değiştirilmiştir.

Farkların varyansının ortalama değerlerin düzeyine göre artış gösterdiği megafon biçimli senaryoya ilişkin farkların analiz sonuçları Tablo 7'de verilmiştir. Sonuçlar incelendiğinde ölçüm teknikleri arasında hem oransal hem de sabit yanlılık bulunmaktadır.

Ölçümlere logaritmik dönüşüm uygulanarak farkların analizi gerçekleştirildiğinde, ölçüm teknikleri arasındaki oransal yanlılık giderilememiş ($r=0,314$, $p<0,001$) aynı zamanda sabit yanlılık da ($t=8,517$, $p<0,001$) tespit edilmiştir (Tablo 7). Oransal yanlılık giderilemediği durumda elde edilecek regresyona dayalı güven sınırları Şekil 3'te elde edilmiştir.

Megafon biçiminde uyarlanmış veri için Tablo 8 sonuçlarında EKK hem sabit hem de oransal yanlılık açısından tutarsız sonuç göstermiş, AEKK yöntemi ise sadece sabit yanlılık açısından tutarsız sonuç vermiştir. Ancak model II regresyon yöntemleri ile hem sabit hem de oransal yanlılık tespit edilmiştir (Tablo 9).

Farklar ve ortalamalar arasında değişen varyanslı ilişki bulunduğunda, Tablo 9 sonuçlarından Ağırlıklı Deming regresyon en dar güven sınırlarını sağlamıştır.

**ŞEKİL 4:** Farklar ve ortalamalar arası regresyona dayalı doğrusal olmayan %95 güven sınırları.

Senaryo 4: Farklar ve Ortalamalar Arasında Eğrisel bir İlişkinin Mevcut Olduğu Durum

Önceki senaryolarda farklar ve ortalamalar arasında doğrusal ilişkiler varsayılarak %95 güven sınırları doğrusal regresyona dayalı olarak oluşturulmuştur. Ancak bu ilişkinin doğrusal olmama durumunda da en iyi eğrisel regresyon denklemine dayalı güven sınırları yöntem kısmında açıklanan adımlar izlenerek ($y= a+ bx+ cx^2$; y: fark, x: ortalama iken) elde edilebilmektedir (Şekil 4).¹⁰

Bu amaçla orijinal veriler saçılım grafiklerinden yararlanarak farklar ve ortalamalar arasında eğrisel bir ilişki elde edilecek şekilde değiştirilmiştir.

Farklar ve ortalamalar arasında eğrisel ilişkili senaryoya ilişkin farkların analiz sonuçları Tablo 10'da verilmiştir. Sonuçlar incelendiğinde ölçüm teknikleri arasında hem oransal hem de sabit yanlılık bulunmaktadır ($p=0,005$, $p=0,002$).

Her iki ölçüme logaritma dönüşümü uygulanarak farklar analiz edildiğinde oransal yanlılık giderilememiş ($r=-0,343$, $p=0,002$), bunun yanı sıra sabit yanlılık da belirlenmiştir (Tablo 11, $t=8,532$, $p=0,002$). Oransal yanlılığın giderilemediği durumda elde edilecek regresyona dayalı güven sınırları Şekil 4'te verilmiştir.

Eğrisel ilişkili model için model I regresyon yöntemi sonuçları Tablo 12'de verilmiştir. Sonuçlar, hem regresyon sabiti hem de regresyon eğimi açısından yine tutarsız bulunmuştur.

Passing-Bablok regresyon yöntemi, ölçümlerin sürekli ölçüğe sahip olmasının yanı sıra doğrusal ilişkili olmasını gerektirmektedir.¹³ Doğrusallık varsayımı Passing Bablok ($p=0,005$) analizi sonucunda reddedilmiştir. Ayrıca katsayılarla ilişkin güven aralıkları elde edilememiştir (Tablo 13).

İki ölçüm tekniği arasında bu şekilde eğrisel ilişki bulunduğu, amaç hataları tespit etmekten ziyade ölçüm tekniklerinin ayarlanması ise Doğrusal Olmayan Deming ve Doğrusal Olmayan Ağırlıklı Deming metodlarının kullanımı önerilir.^{3,12}

Ancak burada amaç, fark ve ortalamalar arasında eğrisel bir ilişki mevcut iken logaritmik dönüşüm ile oransal yanlılık giderilemediğinde; regresyona dayalı olarak %95 güven sınırlarının elde edilebileceğini göstermektir. Bu nedenle doğ-

TABLO 10: İki değerlendiricinin verilerine ait farkların analiz sonuçları.

Orijinal Veri (X - Y) ile (X + Y) / 2 için kurulan Hipotezler	Test istatistiği	Değer	Serbestlik Derecesi	P Değeri
$H_0: \rho = 0$ $H_1: \rho \neq 0$	r	-0,196	200	0,005
$H_0: \bar{D} = 0$ $H_1: \bar{D} \neq 0$	t	7,713	199	0,002
\bar{D} için %95 güven aralığı		2,77 - 4,67		

TABLO 11: İki değerlendiricinin logaritmik dönüşümlü verilerine ait farkların analiz sonuçları.

Log Dönüşümlü Veri log(X-Y) ile log (X+Y) / 2 için kurulan Hipotezler	Test istatistiği	Değer	Serbestlik Derecesi	P Değeri
$H_0: \rho' = 0$ $H_1: \rho' \neq 0$	r	-0,343	200	0,002
$H_0: \bar{D}' = 0$ $H_1: \bar{D}' \neq 0$	t	8,532	199	0,002
X/Y için % 95 güven aralığı		1,069-1,114		

TABLO 12: İki değerlendiricinin verilerine ait model I regresyon yöntemi sonuçları.

Yöntem	a	% 95 GA	b	%95 GA
Denklemler : (1) $E(y)=a+b(x)$ (2) $E(x)=a+b(y)$				
EKK (1) denklemi	8,901	6,66-11,13	0,903	0,86-0,94
EKK (2) denklemi	-4,520	-7,16/ -1,88	1,014	0,97-1,05
AEKK (1) denklemi	7,921	5,98-9,85	0,922	0,872-0,971
AEKK (2) denklemi	-2,034	-4,08-0,014	0,966	0,920-1,012

TABLO 13: İki değerlendiricinin verilerine ait model II regresyon yöntemi sonuçları.

Yöntem	a'	% 95 GA	b'	%95 GA
Deming	6,728	4,760-8,696	0,943	0,907-0,979
Ağırlıklı Deming	-5,031	-7,82/-2,35	1,023	0,976-1,072
Passing- Bablok	0,000	0,000-0,000	1,000	1,000-1,000

rusal olmayan model II regresyon yöntemlerine değinilmemiştir.

TARTIŞMA

Literatürde Deming regresyon yöntemi ya da Bland Altman grafiklerini ele alan birçok çalışma yer almaktadır. Ancak bu çalışmalar arasında farklar ve ortalamalar arası farklı ilişkiler için hem grafik hem de metoda dayalı uyum değerlendirme yöntemlerinin karşılaştırıldığı yayınlar oldukça azdır.

Bir uyum çalışmasındaki amaç, bir ölçüm tekniğini diğerine göre ayarlamak, ya da iki ölçüm tekniği veya değerlendirici arasındaki (sabit ya da oransal) yanlılıkları tespit etmek ise, model II regresyon yöntemleri tercih edilmelidir. Klinik uygulama düzenindeki amaç, özellikle bir ölçüm tekniğinin başka bir ölçüm tekniğini güvenilir şekilde temsil edip edemeyeceğini belirlemek ise Bland & Altman farklar metodu regresyon analizine sadece tamamlayıcı olarak kullanılmalıdır.³

Farklar metodu kullanmaya karar verildiğinde, ilk olarak ortalamalara karşı farkların grafiği çizilmelidir. Farklar, ortalama değerlerin aralığında düzgün (eşit varyanslı) dağılım gösteriyorsa, regresyon analizi gerçekleştirilir. Regresyona ilişkin eğim katsayısı önemsiz bulunmuş ya da korelasyona ilişkin p değeri 0,05'ten büyük ise klasik % 95 uyum sınırları ya da tolerans sınırları hesaplanarak yorumlanabilmektedir.^{1,3,9}

EKK regresyonu için $p < 0,05$ ise, yani farklar ve ortalamalar arasında oransal bir ilişki bulunduğu durumda, x ve y değerlerine logaritma dönüşümü uygulanarak bu ilişki giderilemez ise en uygun çözüm regresyon doğrusu/eğrisi etrafında %95 güven sınırlarını oluşturmaktır.^{6,9}

Çalışmanın analiz sonuçları, Ludbrook (1997), Ludbrook (2010) ve Bland ve Altman (1999) çalış-

malarıyla oldukça uyumludur. Elde edilen bulgulardan görüldüğü gibi model II regresyon yöntemleri sabit ve oransal yanlılığı hem farkların analizi hem de model I regresyon yöntemlerinden daha iyi belirlemiştir.

Bu çalışma, literatürdeki diğer çalışmalardan farklı olarak son zamanlarda çok sık kullanılan uyum karşılaştırmalarında, özellikle ortalama ve farklar arasında oransal yanlılık bulunduğu sırası ile hangi aşamaları izlememiz gerektiği konusuna açıklık getirmektedir. Ayrıca uyum karşılaştırma çalışmalarında regresyona dayalı güven sınırlarının ve diğer regresyon yöntemlerinin eş zamanlı kullanımlarını daha fazla teşvik ederek konu ile ilgili literatüre katkı sağlamaktadır.

SONUÇ

Sonuç olarak yapılan çalışmada, farklar ve ortalamalar arasında oransal bir ilişki bulunduğu; klasik uyum sınırları, uyum sınırlarının güven aralıkları, farkların analizi ve model I regresyon yöntemlerinin tutarlı sonuçlar vermediği gözlenmiştir. Buna karşılık tüm senaryolarda model II regresyon yöntemleri yardımıyla hem sabit hem de oransal yanlılıklar belirlenebilmiştir.

Farklar ve ortalamalar arasında oransal yanlılık bulunmuyorsa klasik %95 uyum ya da tolerans sınırları değerlendirilebilir. Eğer farkların saçılımı ortalama değerlerle birlikte artış gösterirse ya da farklar değişen varyanslı (heteroscedasticity) ise, ayrıca logaritma dönüşümü uygulanmasına rağmen oransal yanlılık giderilemez ise senaryo 2 ve senaryo 3'te görüldüğü gibi ortalamalara göre farkların regresyon doğrusu/eğrisi etrafında regresyona dayalı güven sınırları elde edilebilir. Bu güven sınırları regresyon analizinin varsayımlarıyla uğraşma sorununu ortadan kaldırırken aynı zamanda

iki ölçüm tekniği ile elde edilecek yeni ölçümler arasındaki olası farkı tahmin edebilmeyi sağlar. Söz konusu koşullar için bu çalışma, regresyona dayalı

güven sınırlarının eğrisel ilişkili durumda bile elde edilebileceğini gerçek bir veri seti üzerinde göstermiştir.

KAYNAKLAR

1. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8(2):135-60.
2. Szymtkowski J, Kapala A, Dabrowiecki S. A comparison of statistical methods for the evaluation of diagnostic tests shown on the example of two methods of blood recirculation measurements in dialysis access. *Polish Journal of Surgery* 2009;81(4): 186-92.
3. Ludbrook J. Comparing methods of measurements. *Clin Exp Pharmacol Physiol* 1997; 24(2):193-203.
4. Genç Y, Sertkaya D, Demirtaş S. [Statistical methods for assessing agreement between two measurement technique in clinical research]. *Ankara Üniversitesi Tıp Fakültesi Mecmuası* 2003;56(1):1-6.
5. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476): 307-10.
6. Ludbrook J. Confidence in Altman-Bland plots: a critical review of the method of differences. *Clin Exp Pharmacol Physiol* 2010;37(2):143-9.
7. Dewitte K, Fierens C, Stöckl D, Thienpont LM. Application of the Bland-Altman plot for interpretation of method-comparison studies: a critical investigation of its practice. *Clin Chem* 2002;48(5):799-801.
8. Euser AM, Dekker FW, le Cessie S. A practical approach to Bland-Altman plots and variation coefficients for log transformed variables. *J Clin Epidemiol* 2008;61(10):978-82.
9. Altman DG, Gardner MJ. Regression and correlation. In: Altman DG, Machin D, Bryant TN, Gardner MJ, eds. *Statistics with Confidence*. 2nd ed. London: BMJ Books; 2000. p.73-92.
10. Linnet K. Evaluation of regression procedures for methods comparison studies. *Clin Chem* 1993;39(3):424-32.
11. Saraçlı S, Yılmaz V, Doğan İ. [Simple linear regression techniques in measurement error models]. *Anadolu University Journal of Science and Technology* 2009;10(2):335-42.
12. Saraçlı S, Doğan İ, Doğan N. [Deming regression technique in medical method comparison studies]. *Türkiye Klinikleri J Biostat* 2009;1(1):9-15.
13. Biliç-Zulle L. Comparison of methods: Passing and Bablok regression. *Biochem Med (Zagreb)* 2011;21(1):49-52.