

UZUNLAMASINA VERİLERİN YER ALDIĞI FAKTÖRİYEL TASARIMLAR İÇİN PARAMETRİK OLMAYAN YAKLAŞIM

Can ATEŞ* Yasemin GENÇ** Meliha HAFIZ*** Derya ÖZTUNA****
Şehim KUTLAY***** Bülent SEÇKİN*****

ÖZET

Uzunlamasına verilerin analizinde parametrik yöntemler oldukça yaygın olarak kullanılsa da bu yöntemlerin kullanımı, belirli varsayımların sağlanmasına bağlıdır. Son yıllarda parametrik test varsayımlarının sağlanmadığı verilerde ve/veya sıralı kategorik değişkenler kullanılarak yapılan değerlendirmelerde birden çok faktörün etkisini incelemek üzere parametrik olmayan yöntemler geliştirilmiştir. Bu çalışmanın amacı, uzunlamasına verilerin analizinde Brunner ve arkadaşları tarafından geliştirilen parametrik olmayan yöntemin tanıtılması ve Fiziksel Tıp ve Rehabilitasyon alanından bir veri seti üzerinde uygulamasının yapılmasıdır. Uzunlamasına verilerin parametrik olmayan yöntemlerle analizinde, parametre tahminlerinin elde edilmesinde ve test edilmek istenen hipotezlerin oluşturulmasında marjinal dağılımlar kullanılır. Hipotezlerin test edilmesinde örneklem büyüklüğüne bağlı olarak ANOVA tipi veya Wald tipi istatistik kullanılırken; tanımlayıcı istatistik olarak göreceli deneme/marjinal etkileri yorumlanır. Uygulama için tedavi öncesi, tedavi sırasında ve tedaviden 3 ay sonra elde edilen Duruöz El İndeksi değerleri bağımsız iki grupta (tedavi ve kontrol) F1-LD-F1 tasarımı kullanılarak karşılaştırılmıştır. Brunner ve arkadaşları tarafından geliştirilen yöntem, uzunlamasına verilerin analizinde ortaya çıkabilecek farklı deneme tasarımlarına çözüm sunmaktadır. Herhangi bir test varsayımı gerektirmemesi ve örneklem büyüklüğü bakımından kısıtı olmaması yöntemin avantajları olarak özetlenebilir.

Anahtar Kelimeler: Parametrik olmayan yöntemler, Tekrarlı ölçümler, Uzunlamasına çalışmalar.

1. GİRİŞ

Uygulamalı bilim dallarında yapılan araştırmalarda, tekrarlı ölçüm içeren çalışma tasarımları sıklıkla kullanılır. Tekrarlı ölçümler, aynı denekten farklı zaman noktalarında ya da farklı koşullarda gözlemlerin alınması ile oluşur. Tekrarlı ölçümlerin zamana bağlı olduğu durumda elde edilen veriler, literatürde yaygın olarak “uzunlamasına/izlemsel veriler (longitudinal data)” olarak adlandırılır.

*Uzman Dr., Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik AD, Ankara, e-posta can.ates@gmail.com

**Doç. Dr., Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik AD, Ankara, e-posta genc@medicine.ankara.edu.tr

***Uzman Dr., Medical Park Fizik Tedavi ve Rehabilitasyon Bölümü, Gaziantep, e-posta melihat@yahoo.com

****Doç. Dr., Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik AD, Ankara, e-posta dgokmen2001@yahoo.com

*****Prof. Dr., Ankara Üniversitesi, Tıp Fakültesi, Fiziksel Tıp ve Rehabilitasyon AD, Ankara, e-posta skutlay@medicine.ankara.edu.tr

*****Prof. Dr., Ankara Üniversitesi, Tıp Fakültesi, Fiziksel Tıp ve Rehabilitasyon AD, Ankara, e-posta seckin@medicine.ankara.edu.tr

Uzunlamasına verilerin kullanıldığı çalışmalarda, zamana bağlı ölçümlerin yanı sıra bağımsız grup faktörü/faktörleri de olabilir. Grup faktörünün de bulunduğu durumlarda temel amaç, incelenen değişkenin denemeler arasındaki farklılığının (grup faktörü), zaman içerisindeki değişiminin (zaman faktörü) ve zaman içerisindeki değişiminin gruplarda benzer olup olmadığının (grup x zaman etkileşimi) incelenmesidir. Uzunlamasına verilerin en doğru biçimde çözümlenmesi ve yorumlanması ile ilgili olarak yapılan çalışmalar son yıllarda artış göstermiş ve bu verilerin analizinde kullanılmak üzere parametrik ve parametrik olmayan yöntemler geliştirilmiştir.

Uzunlamasına verilerin analizinde parametrik yöntemler oldukça yaygın olarak kullanılsa da bu yöntemlerin kullanımı, belirli varsayımların (normallik varsayımı, varyansların homojenliği vs.) sağlanmasına bağlıdır. Bağımsız grupların ve zamana bağlı ölçümlerin olduğu bir çalışmada, parametrik test varsayımlarının sağlanmadığı ve grup x zaman etkileşiminin anlamlı bulunduğu durumda, zaman içindeki değişim her bir bağımsız grup için ayrı ayrı ele alınacağından veri seti tek seferde analiz edilememektedir. Veri üzerinde bu tür bir çözümleme, çalışmanın başında planlanan Tip I hata miktarını arttıracaktır (Abdi, 2007). Bunun yanında sağlık alanında yapılan çalışmalarda sıralı kategorik (ordinal) veri yapısına da sıklıkla rastlanılmaktadır. Özellikle kategori sayısının az olduğu veri yapılarında parametrik test varsayımlarının sağlanması mümkün olmamaktadır ve bu durumlarda kullanılmak üzere, parametrik olmayan yöntemler geliştirilmiştir (Hollander vd., 1974).

Uzunlamasına verilerin analizinde kullanılan parametrik olmayan yöntemlerin ilk kullanım alanı eşleştirilmiş (bağımlı) iki örneklem tasarımıdır. Bu tasarımda test edilen hipotez, her iki örneklemin marjinal dağılımları kullanılarak formüle edilmiştir. Marjinal dağılımlar, incelenen değişkenlerin zamana bağlı değişim gösterip göstermediğinin araştırılmasında kullanılır. Bu durumda “zamana bağlı değişim yoktur” hipotezi aynı zamanda “marjinal dağılımlar eşittir” hipotezidir. Brunner ve Neumann, 1982 yılında aynı denekten birden fazla ölçümün alındığı deney düzenleri için sıra sayısı (rank) metodlarını önermiştir. Thompson (1990, 1991), Brunner ve Neumann (1982)’nin sonuçlarını genelleştirmiş ve doğrusal sıra sayısı istatistiklerini basit deneme düzenlerinde kullanmıştır. Bu çalışmalarda, marjinal dağılımlarla ilgili tek varsayım süreklilik varsayımıdır. “Normalleştirilmiş versiyon” olarak adlandırılan dağılım fonksiyonu kullanıldığında, sürekli dağılım fonksiyonu varsayımına gerek duyulmayabilir (Kruskal, 1952; Ruymgaart, 1980). Bu basit yaklaşım, kesikli verilerin kısmen de sıralı kategorik verilerin analizine olanak sağlar.

Bu çalışmanın amacı, uzunlamasına verilerin analizinde Brunner ve arkadaşları tarafından geliştirilen parametrik olmayan yöntemi tanıtmak ve sağlık alanında, karpal tünel sendromu tanısı almış bir grup hastadan elde edilen veri seti üzerinde bu yöntemin uygulamasını yapmaktır.

2. YÖNTEM

Uzunlamasına verilerin olduğu bir deney tasarımında, $i=1, \dots, a$ grupları, $k=1, \dots, n_i$ her bir gruptaki bağımsız deney birimini ve $s=1, \dots, t$ ölçüm alınan farklı noktaları ve son olarak, X_{ik1}, \dots, X_{ikt} ölçüm değerlerini gösterebilir (Tablo 1). i . gruptaki k . birimin t sayıdaki ölçümü vektör formunda aşağıdaki gibi yazılabilir:

$$X_{ik} = \begin{pmatrix} X_{ik1} \\ \vdots \\ X_{ikt} \end{pmatrix} \quad (1)$$

Bu deney tasarımında X_{ik} vektörleri bağımsız iken; grubun her bir düzeyinde her bir denek için elde edilen zamana bağlı ölçümlerin oluşturduğu vektörler birbirine bağımlıdır (Brunner vd., 2002).

Tablo 1. Temel gösterimler (Brunner vd., 2002)

Faktör A	Denekler	Zaman		
		$s = 1$...	$s = t$
$i = 1$	$k = 1$	X_{111}	...	X_{11t}
	\vdots	\vdots	\vdots	\vdots
\vdots	$k = n_1$	X_{1n_11}	...	X_{1n_1t}
	\vdots	\vdots	\vdots	\vdots
$i = a$	$k = 1$	X_{a11}	...	X_{a1t}
	\vdots	\vdots	\vdots	\vdots
	$k = n_a$	X_{an_a1}	...	X_{an_at}

X_{ik} , $F_i(x)$ normalleştirilmiş dağılım fonksiyonuna sahip rastgele bir değişkendir. Normalleştirilmiş dağılım hem sürekli hem de kesikli dağılımlarda kullanılabilir özelliğine sahiptir. Bir gruptaki farklı deneklerden elde edilen gözlemler, bir denemenin tekrarı olarak düşünülebileceğinden, X_{ik} vektörlerinin $G_i(x)$ ortak dağılım fonksiyonuna sahip olduğunu varsaymak mümkündür. Yani dağılım fonksiyonları k indisine bağlı değildir.

X_{ik} vektörünün t tane marjinal dağılım fonksiyon vektörü,

$$F_i = (F_{i1}, \dots, F_{it})'$$

olmak üzere, a . grupta, t . zaman noktasındaki tüm marjinal dağılımlar vektörü,

$$F = (F_1', \dots, F_a') = (F_{11}, \dots, F_{1t}, \dots, F_{a1}, \dots, F_{at})'$$

olarak gösterilir. Deneme etkilerini tanımlamak ve ilgili hipotezleri kurmak için marjinal dağılımların kullanıldığı modele 'parametrik olmayan marjinal model' adı verilir. Bu modelin altında yatan temel düşünce, gözlemlerin bağımsızlık yapısının kullanılması ve rastgele vektörlerin marjinal dağılımlarının aynı olup olmadığının araştırılmasıdır (Brunner vd., 2002).

2.1 Tasarımlar

Uzunlamasına verilerin parametrik olmayan yöntemlerle analizinde, çalışmada bağımsız grup faktörünün ve alt zaman faktörünün olup olmamasına göre farklı deneme tasarımları mevcuttur.

LD-F1 Tasarımı: Uzunlamasına veriler için en basit deneme düzeni olan bu tasarımda, homojen bir denek grubunda her bir denekten, t zaman noktasında ölçüm alınır.

LD-F2 Tasarımı: Bu tasarımda, homojen bir denek grubunda, her bir bireyden c zamanda ve her bir durum için t kez ölçüm alınır.

F1-LD-F1 Tasarımı: Uzunlamasına verilerde en sık kullanılan bu tasarımda, her biri n_i denekten oluşan a gruptan t zaman noktasında ölçüm alınır.

F2-LD-F1 Tasarımı: Bu tasarımda, a grupta (A faktörü) yer alan ve b alt grupta (B Faktörü) tabakalandırılmış deneklerin her birinden t zaman noktasında ölçüm alınır.

F1-LD-F2 Tasarımı: Bu tasarımda her biri n_i bireyden oluşan a grupta (A faktörü), her birey c zamanda (C faktörü) ve her bir durum için t (T faktörü) kez gözlenir.

2.2 Göreli Etkiler

Parametrik olmayan modelde,

$$X_{ik} = (X_{ik1}, \dots, X_{ikt})'$$

vektörlerinden elde edilen rastgele değişkenler,

$$Y_{ik} = \varphi(X_{ik1}, \dots, X_{ikt})$$

ile gösterilir. Uzunlamasına verilerin analizinde ilgilenilen durum, her zaman noktasına ait farklılıktan çok, toplam farklılığın araştırılması olduğundan, her denekten her bir zaman noktasında alınan gözlemler için özet fonksiyon $\varphi(X_{ik1}, \dots, X_{ikt})'$ ile tanımlanır (Brunner ve Puri, 2001; Brunner vd., 2002). Y_{ik} değerleri üzerindeki tek varsayım,

$$\text{bağımsız ve } G_i(y) = \frac{1}{2} [G_i^+(y) + G_i^-(y)]$$

dağılım fonksiyonu ile aynı dağılıma sahip olmalarıdır. i . deneme için “görelî özet etki (p_i)”, deney tasarımındaki tüm ($i = 1, \dots, a$) dağılımına göre G_i yardımıyla aşağıdaki gibi tanımlanır (Brunner ve Puri, 2001; Brunner vd., 2002; Brunner ve Puri, 2002):

$$p_i = \int H(y) dG_i(y) = \int H dG_i \quad (2)$$

Bu eşitlikte,

$$H(y) = 1/N \sum_{i=1}^a n_i G_i(y),$$

her biri n_i büyüklüğünde $i=1, \dots, a$ grup için özet değişkenlerin dağılım fonksiyonlarının ağırlıklı ortalaması ve

$$N = \sum_{i=1}^a n_i$$

çalışmada kullanılan toplam denek sayısıdır. Eşitlik 2'den de görülebileceği gibi, her bir grubun dağılım fonksiyonu o grupta yer alan gözlem sayısı ile çarpıldığından, daha fazla sayıda gözlem içeren gruba daha fazla ağırlık verilmiş olur.

H dağılım fonksiyonu ile Z rastgele değişkeni, G_i dağılım fonksiyonu ile Y_i rastgele değişkeninden bağımsız ise, p_i görelî etkisi aşağıdaki gibi tanımlanır (Brunner ve Puri, 2001):

$$p_i = P(Z < Y_i) + \frac{1}{2} P(Z = Y_i) \quad (3)$$

Bu durumda, görelî etkinin $\frac{1}{2}$ 'den küçük (büyük) olması, Y_i ' den elde edilen gözlemlerin, Z rastgele değişkeninden elde edilenlerle kıyaslandığında, daha küçük

(büyük) değerlere sahip olma eğilimi göstereceğini ifade eder. Bu durum, bir gruptaki gözlemlerin, H dağılımına sahip gözlemlerle ilişkisini tanımlayan görelî etki tanımındaki, “görelî” sıfatını daha net açıklar (Brunner ve Puri, 2001; Brunner vd., 2002).

Sonuç olarak, $G = (G_1, \dots, G_a)'$ dağılım fonksiyonları vektöründen yararlanılarak, görelî özet etkiler vektörü aşağıdaki gibi tanımlanır:

$$p = \int H dG = (p_1, \dots, p_a)' \quad (4)$$

2.3 Görelî Marjinal Etkiler

Parametrik olmayan marjinal modellerde, deneme etkileri toplam marjinal dağılımlar kullanılarak tanımlanır. Örneğin, F1-LD-F1 tasarımı için toplam marjinal dağılım, Eşitlik 5'te gösterilmiştir:

$$H = 1/N \sum_{i=1}^a \sum_{s=1}^t n_i F_{is} \quad (5)$$

Bu tanımdan yola çıkarak, i . grubun s . zaman noktasında H fonksiyonuna göre görelî marjinal etkisi,

$$p_{is} = \int H dF_{is}$$

olarak tanımlanır. Yani p_{is} , F_{is} marjinal dağılımının H 'ye göre eğilimini gösterir. F_{is} , H 'nin solunda kalan alanda değerler almaya eğilimli ise $p_{is} < 1/2$ ' dir. F_{is} , H 'nin sağında kalan alanda değerler almaya eğilimli ise $p_{is} > 1/2$ ' dir. F_{is} , için herhangi bir eğilim olmadığı durumda ise $p_{is} = 1/2$ ' dir (Brunner ve Puri, 2001; Shah ve Madden, 2003; Gao vd., 2005). Yine F1-LD-F1 tasarımından elde edilen, tüm ($a \times t$) görelî marjinal etkiler, vektör formunda aşağıdaki gibi ifade edilebilir;

$$\begin{aligned} p &= (p_{11}, \dots, p_{1t}, \dots, p_{a1}, \dots, p_{at})' \\ &= (\int H dF_{11}, \dots, \int H dF_{1t}, \dots, \int H dF_{a1}, \dots, \int H dF_{at})' \\ &= \int H dF \end{aligned} \quad (6)$$

2.4 Hipotez Testleri

Parametrik olmayan yöntemlerde hipotezler, dağılım fonksiyonları ya da görelî etkiler yardımıyla oluşturulur. Modelde yer alan ana etkiler için hipotezler, Y_{ik} rastgele değişkenlerinin dağılım fonksiyonları (G_i) ya da görelî özet etkileri (p_i) kullanılarak, etkileşim terimleri ise marjinal dağılımlar (F_{is}) ya da görelî marjinal etkiler (p_{is}) kullanılarak formülize edilir (Brunner ve Puri, 2001; Brunner vd., 2002).

2.4.1 Test İstatistikleri

Parametrik olmayan yöntemlerde, hipotezlerin test edilmesinde üç farklı test istatistiğinden yararlanılır. Bunlar; Wald Tipi İstatistik (WTİ), ANOVA Tipi İstatistik (ATİ) ve Hotelling T^2 istatistiğidir.

2.4.1.1. Wald tipi istatistik (WTİ)

C kontrast matrisi olmak üzere, $H_0^F: CF = \mathbf{0}$ hipotezi altında,

$$\sqrt{n}C\hat{p} = \sqrt{n}C\frac{1}{n}\bar{R} = \frac{1}{\sqrt{n}}C(\bar{R}_1, \dots, \bar{R}_a)'$$

bağıntı vektörü, beklenen değer vektörü $\mathbf{0}$ ve kovaryans matrisi CV_nC' ile asimptotik olarak çok değişkenli normal dağılım gösterir (Ruymgaart, 1980; Mynzel ve Brunner, 2000).

$$\hat{V}_i = \frac{1}{N^2(n_i-1)}\sum_{k=1}^{n_i}(R_{ik} - \bar{R}_i)(R_{ik} - \bar{R}_i)'$$

olmak üzere, V_n kovaryans matrisinin tutarlı tahmin edicisi,

$$\hat{V}_n = \bigoplus_{i=1}^a \frac{n}{n_i} \hat{V}_i \text{ dir.}$$

$H_0^F: CF = \mathbf{0}$ hipotezi altında,

$$Q_n(C) = n\hat{p}'C'[C\hat{V}_n C']^{-1}C\hat{p} \quad (7)$$

Karesel (quadratic) formu, V_n 'in düzenli (regular) matris olması koşuluyla, asimptotik olarak $f = \text{rank}(C)$ serbestlik derecesi ile merkezi ki kare dağılımına (χ_f^2) sahiptir. $Q_n(C)$ istatistiği, WTİ'nin sıra numarası versiyonu olarak adlandırılır (Brunner ve Puri, 2001; Brunner vd., 2002). Bu test istatistiği, büyük örneklem genişliklerinde ($n > 200$) daha güvenilir sonuçlar sağlarken, küçük ya da orta büyüklükteki örneklerde kullanımı önerilmez (Ruymgaart, 1980).

2.4.1.2 Anova tipi istatistik (ATİ)

WTİ'de temel problem, bilinmeyen V_n kovaryans matrisinin tutarlı bir tahmin edici ile yer değiştirilmesi gerekliliğidir. Tahmin edilen her varyans ve kovaryans, karesel formun değişkenliğine katkıda bulunur. Bu durumda en basit çözüm, eşitlik 7'deki \hat{V}_n kovaryans matrisinin çıkarılması ve aşağıdaki istatistiğin kullanılmasıdır (Akritas ve Brunner, 1997; Munzel ve Puri, 2000; Brunner ve Puri, 2001; Brunner vd., 2002).

$$Q_n^*(C) = n\hat{p}'C'[CC']^{-1}C\hat{p}. \quad (8)$$

2.4.1.3 Hotelling T² istatistiği

Bu istatistiğe ait açıklamalar, homojen tek bir grubun olduğu deneme tasarımları ile kısıtlanmıştır (LD-F1, LD-F2). İlk olarak, bağımsız t -boyutlu normal dağılım gösteren rastgele vektörler $Z_k \sim N(\mu, S)$ ($k=1, \dots, n$) olsun.

$$\bar{Z} = \frac{1}{n}\sum_{k=1}^n Z_k$$

ortalamayı,

$$\hat{S}_n = \frac{1}{n-1}\sum_{k=1}^n (Z_k - \bar{Z}.)'(Z_k - \bar{Z}.)'$$

ise, bu bağımsız değişkenlere ilişkin örneklemeden elde edilen kovaryans matrisini gösterir. μ 'lerin tüm bileşenlerinin eşit olduğu hipotezi,

$$H_0^{\mu}: \mu_1 = \dots = \mu_t,$$

matris gösterimiyle; P_t t boyutlu merkez matrisi olmak üzere,

$$H_0^{\mu}: P_t \mu = 0$$

olarak ifade edilir. Bu ifadeler kullanılarak elde edilen Hotelling T^2 istatistiği (Brunner vd., 2002).

$$T^2 = n\bar{Z}' \cdot P_t (P_t \hat{S}_n P_t)^{-1} P_t \bar{Z}. \quad (9)$$

olup; H_0^{μ} hipotezi altında, asimptotik olarak merkezi ki kare dağılımına (χ_{t-1}^2) sahiptir.

3. BULGULAR

3.1 Hastaların Seçimi ve Çalışma Protokolü

Çalışma kapsamında yapılan uygulamalar, Ankara Üniversitesi Tıp Fakültesi (AÜTF) Fiziksel Tıp ve Rehabilitasyon (FTR) Anabilim Dalında gerçekleştirilen bir tez çalışması için elde edilen veri seti üzerinde gerçekleştirilmiştir (Hafız, 2010). Bu çalışma Helsinki Deklarasyonu 2008 prensiplerine uygun olarak yürütülmüş olup; etik kurul onayı alınmıştır (Etik kurul kabul no: 154-4937, 29.06.2009).

Bu tez çalışmasında 2007-2009 yılları arasında AÜTF, FTR Anabilim Dalı Genel Polikliniğine başvuran, klinik bulgular ve elektronöromiyografi (ENMG) sonuçları ile karpal tünel sendromu (KTS) tanısı alan toplam 59 birey alınmıştır. Öncelikle çalışmaya alınan tüm bireylere standart KTS tedavisi [el bileği ön yüzüne 10 dakikalık yüzeyel ısı (Hotpack (HP) - sıcak uygulama) tedavisi] uygulanmıştır. Tedavi grubuna standart KTS tedavisine ek olarak 4 hafta süre ile el bileği ön yüzüne 5 dakika süreyle 1W/cm² dozda ultrason (US) ile %1'lik hidrokortizon pomat (fonoforez) uygulanırken, kontrol grubuna ise fonoforez yapılmadan aynı süre ve dozda US tedavisi uygulanmıştır (Hafız, 2010). Çalışmaya alınan bireylerin gruplara rastgele atanmasında (randomizasyon) Random Allocation Software programı kullanılarak, blok randomizasyon yöntemi ile elde edilen atama şeması takip edilmiştir. Çalışma sonucunda analiz aşamasında tedavi grubunda 31, kontrol grubunda 28 KTS hastası bulunmaktadır.

Uygulanan tedavilerin, hastaların el ve el bileği aktiviteleri üzerindeki etkilerinin değerlendirilmesi amacıyla klinik ve analitik olmak üzere birden fazla test uygulanmış, bu çalışma kapsamında bu analitik testlerden Duruöz El İndeksi (DEİ) kullanılmıştır. 1996 yılında romatoid el için spesifik olarak üretilmiş olan bu indeks, 18 el ve el bileği aktivitesi içeren fonksiyonel yetersizlik göstergesinden oluşmaktadır. Yanıt kategorileri 6'lı Likert [0 (Hiç zorluk çekmiyorum) – 5 (Çok zorlukla)] türünde olup; toplam skor 0-90 arasında değişmektedir. Skor yüksekliği aktivite yetersizliğinin fazla olduğunu göstermektedir (Duruöz vd., 1996).

Tedavi öncesi, tedavi sırasında ve tedaviden 3 ay sonra elde edilen DEİ değerlerinin zamana bağlı tedavi ve kontrol gruplarında incelendiği bu çalışmada F1-LD-F1 tasarımı

kullanılmıştır. Bu veri setinde, tedavi ve kontrol olmak üzere bağımsız iki grup (Faktör A) ve zamana bağlı üç tekrar (Faktör T) bulunmaktadır (Ateş, 2010).

3.2 Hipotezler

DEİ' n den elde edilen skorlardan, ortalama grup etkisi;

$$H_0^F(A): F_{1.} = F_{2.},$$

zaman etkisi;

$$H_0^F(T): F_{.1} = F_{.2} = F_{.3}$$

grup x zaman etkileşimi etkisi;

$$H_0^F(AT): F_{11} = F_{12} = F_{13} = F_{21} = F_{22} = F_{23}$$

yokluk hipotezleri ile test edilir. Test istatistikleri Tablo 2'de, görelî marjinal etkiler Tablo 3'te özetlenmiştir.

Tablo 2. DEİ skorlarından elde edilen test istatistikleri

	Wald (WTİ)	Tipi	İstatistikler	ANOVA Tipi	İstatistikler (ATİ)
FAKTÖRLER		Q_n	p	F_n	p
A (Grup)		5.916	0.015	5.916	0.018
T (Zaman)		137.350	<0.001	68.383	<0.001
AT (Grup x Zaman)		34.122	<0.001	17.392	<0.001

Çalışmanın örneklem büyüklüğü 200'den küçük olduğu için hipotezlerin test edilmesinde ATİ kullanılmış ve grup, zaman ve etkileşim terimleri istatistiksel olarak anlamlı bulunmuştur (sırasıyla; $p=0.018$, $p<0.001$, $p<0.001$) (Tablo 2).

Grup x zaman etkileşiminin anlamlı bulunması, zaman noktaları arası değişimin iki grupta benzer olmadığı anlamına gelir. Bu nedenle doğru yorumların yapılabilmesi için, iki grubun farklı zaman noktalarındaki görelî marjinal etkileri, bu etkilere ait %95 güven aralıkları Tablo 3'te sunulmuştur. Çalışmamızda yer alan örnek için iki grup ve üç zaman noktası bulunduğundan, toplam altı adet görelî marjinal etkinin ortalamasının 0.50 olması beklenir. Bu bilgiler ışığında, tedavi grubunun tedavi öncesi, tedavi sırasında ve tedaviden 3 ay sonraki görelî marjinal etki değerleri sırasıyla 0.618, 0.381 ve 0.276 iken, aynı değerler kontrol grubunda 0.644, 0.571 ve 0.530 olarak elde edilmiştir. Anlaşılacağı gibi, her iki grupta da verilen tedaviler ile DEİ skorlarında bir düşüş gözlenmiş, ancak tedavi grubundaki düşüş, kontrol grubundakine kıyasla daha fazla olmuştur. (Tablo 3). Görelî marjinal etkiler ve onlara ilişkin %95 güven

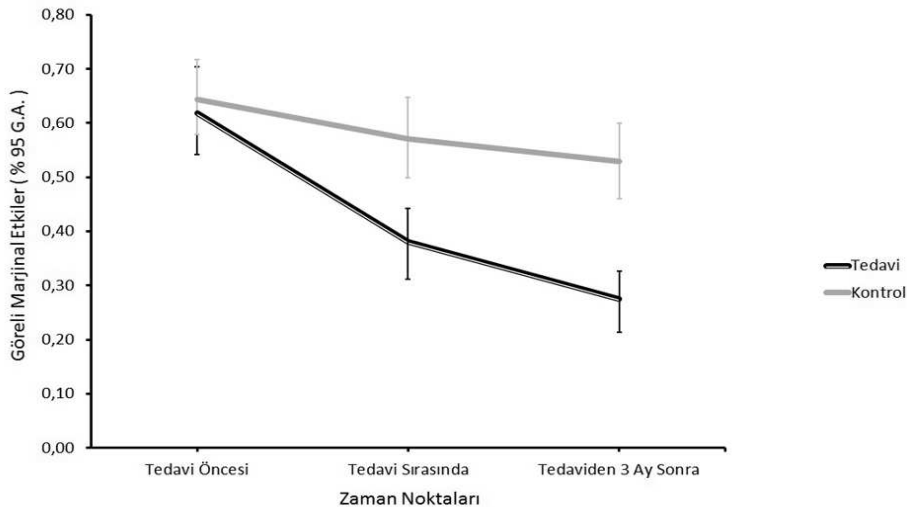
aralıklarından elde edilen grafikte, gruplardaki zamana bağlı değişimleri görsel olarak daha net ifade etmektedir (Şekil 1).

Tablo 3. DEİ skorları için farklı grup ve zaman noktalarına ait görece marjinal etkiler ve %95 güven aralıkları

Grup	Zaman Noktaları (Görelî Marjinal Etkiler (%95 G.A.))		
	Tedavi Öncesi	Tedavi Sırasında	Tedaviden 3 Ay Sonra
Tedavi	0.618 (0.532-0.694)	0.381 (0.319-0.450)	0.276 (0.226-0.339)
Kontrol	0.644 (0.570-0.709)	0.571 (0.494-0.643)	0.530 (0.460-0.599)

Tedavi ve kontrol grupları için tedavi öncesinde elde edilen DEİ skorlarına ilişkin görece marjinal etkiler birbirlerine yakın bulunmuştur. Tedavi sırasındaki görece marjinal etkileri her iki grupta da tedavi öncesine göre düşüş göstermekle beraber; tedavi grubunda daha belirgindir. Benzer yorum, tedaviden 3 ay sonraki ile tedavi sırasındaki görece etkiler arası farklılık için de geçerlidir. Veri setimizde kayıp gözlem değerleri olması nedeniyle, zaman noktaları arasında ikili karşılaştırmalara ilişkin p değerleri elde edilememiş farklılığı yaratan zaman noktalarının tespit edilmesi amacıyla marjinal etkilere ilişkin %95 güven aralıkları kullanılmıştır.

Tedavi öncesi değerlerinden iki grup için hesaplanan güven aralıklarının birbiriyle çakışması, tedavi ve kontrol grupları arasında farklılık olmadığını; ikinci ve üçüncü zaman noktalarında ise birbirlerinden tamamen ayrılan güven aralıklarının iki grubun istatistiksel olarak farklılığını ortaya koymaktadır. Sonuç olarak her iki grupta da incelenen zaman boyunca DEİ skorlarında bir düşüş gözlenmekte ancak bu düşüşler, tedavi grubunda anlamlı olarak daha belirgin ortaya çıkmaktadır.



Şekil 1. İki grupta DEİ skorlarına ait marjinal görece etkilerin ve %95 güven aralıklarının zamana bağlı değişimi

4. TARTIŞMA VE SONUÇ

Sağlık alanındaki araştırmalarda aynı denekten farklı zaman noktalarında ya da farklı koşullarda gözlemlerin alınması ile oluşan tekrarlı ölçüm içeren çalışma tasarımları sıklıkla kullanılır. Tekrarlı ölçümlerin zamana bağlı olduğu durumda elde edilen veriler, literatürde yaygın olarak “uzunlamasına/izlemsel veriler (longitudinal data)” olarak adlandırılır. Uzunlamasına verilerin kullanıldığı çalışmalarda, zamana bağlı ölçümlerin yanı sıra bağımsız grup faktörü/faktörleri de olabilir.

Uzunlamasına verilerin analizinde parametrik yöntemler oldukça yaygın olarak kullanılsa da bu yöntemlerin kullanımı, belirli varsayımların sağlanmasına bağlıdır. Ancak, sıralı kategorik (ordinal) veri yapılarının kullanıldığı çalışmalarda, özellikle de kategori sayısının az olduğu durumlarda, bu varsayımların sağlanması mümkün olmamaktadır (Munzel ve Bandelow, 1998). Ayrıca birden çok faktörün etkisinin incelendiği çalışmalarda, parametrik olmayan klasik yöntemler kullanılarak yapılan analizlerde, veri setinin tek seferde analiz edilmesi mümkün olmamaktadır. Bu nedenlerle, son yıllarda bu tür araştırma tasarımlarında kullanılmak üzere parametrik olmayan yöntemler geliştirilmiştir.

Brunner ve Neumann, 1982 yılında aynı denekten birden fazla ölçümün alındığı deney düzenleri için sıra sayısı (rank) metodlarını önermiş, daha sonra Thompson (1990, 1991), bu yöntemleri genelleştirerek, doğrusal sıra sayısı istatistiklerini basit deneme düzenlerinde kullanmıştır. Bu düzenlerde, yokluk hipotezleri incelenen faktörlerin marjinal dağılımları esas alınarak oluşturulurken, bu hipotezlerin istatistiksel önemliliklerini test edilmesinde örneklem büyüklüğüne bağlı olarak farklı test istatistikleri kullanılır. Elde edilen sonuçların özetlenmesinde ise, tanımlayıcı istatistik olarak görece deneme etkileri ve/veya görece marjinal etkiler yorumlanır.

Son yıllarda sağlık alanında yapılan çalışmalarda da yöntemle ilişkin örneklerle sıklıkla rastlamak mümkündür. Gao, 2004 yılında yüksek lisans tezi olarak kabul edilen çalışmasında rastgele 3 gruba ayrılan (düşük doz, yüksek doz ve plasebo) toplam 120 hastadan başlangıç, 1. ve 3. ayların sonunda olmak üzere 3 zaman noktasında kan ve idrar örnekleri almış yeşil çayın içerdiği polifenolün kan ve idrar örneklerinde aflatoksin konsantrasyonlarını nasıl etkilediğini incelemiş, aynı çalışmada F1-LD-F1 ve LD-F1 tasarımlarının kullanımıyla elde edilen sonuçları yorumlamıştır.

Benzer şekilde, Albus ve arkadaşları tarafından 2009 yılında gerçekleştirilen randomize kontrollü çalışmada, yine F1_LD_F1 tasarımı kullanılmıştır. Bağımsız iki gruba (kontrol ve müdehale) oluşturan toplam 76 hasta, 7 sene içinde 4 zaman noktasında kardiyovasküler hastalık taramasından geçirilmişler ve sonuçta sadece grupların zamana göre değişimini ifade eden etkileşim teriminde anlamlı farklılık bulunmuştur.

Diş sağlığı ve tedavisi alanında, Kamburoglu ve arkadaşları da, 2010 yılında yaptıkları çalışmada, dişin beş farklı bölgesini, 4 ayrı görüntüleme yöntemiyle çürük seviyeleri bakımından karşılaştırmayı amaçlamışlar ve bunun için F1_LD_F1 tasarımı kullanmışlardır. Sonuç olarak ise, bölgeler, görüntüleme yöntemleri ve bu iki faktörün etkileşim terimi bakımından anlamlı farklılık bulmuşlardır.

Bu çalışma kapsamında ise, KTS'nin tedavisinde US ve fonoforez tedavisine karşılık sadece US tedavisinin etkinliğinin incelendiği veri seti, uzunlamasına verilerin parametrik olmayan yaklaşımlarla analizinde en yaygın kullanılan tasarımlardan biri

olan F1-LD-F1 tasarımı kullanılarak değerlendirilmiştir. Hastaların el ve el bileği aktivitelerinin DEİ ile incelendiği bu veri setinde, tedavi ve kontrol olmak üzere bağımsız iki grupta (Faktör A), hastaların tedavi sürecindeki değişimleri, tedavi öncesi, tedavi sırasında ve tedaviden 3 ay sonra olmak üzere 3 zaman noktasında (Faktör T) izlenmiştir. Faktörlere ilişkin ana etkilerin ve etkileşim teriminin anlamlılığı ATİ ile değerlendirildiğinde, her üç etki de istatistiksel olarak anlamlı bulunmuştur. Zamana bağlı değişimlerin gruplarda farklılık gösterdiğinin belirlendiği bu çalışmada US ve fonofrez tedavisinin uygulandığı gruptaki hastalarda gözlenen iyileşme (DEİ skorlarındaki düşüş), sadece US tedavisinin verildiği gruptakilere kıyasla daha fazladır. Bu sonuçlar KTS'nin tedavisinde kombine US-fonofrez tedavisinin, tek başına US tedavisine göre daha etkin olduğunu göstermektedir.

Çalışmada, denek sayısı göz önünde bulundurularak, ATİ kullanılmıştır (Brunner ve Puri, 2002; Shah ve Madden, 2003). Munzel ve Brunner, 2000 yılında yaptıkları bir benzetim çalışmasında, farklı dağılımlardan dört tekrarlı ölçüm ve 3 bağımsız gruptan oluşan veri setleri türetmiştir. Gruplara düşen denek sayısının 7, 20 ve 50 olduğu bu çalışmada ATİ ve WTİ'nin performanslarını kıyaslamışlardır. Sonuçta ATİ'nin en küçük örneklem çapında dahi güçlü sonuçlar verdiğini, WTİ'nin ise gruplara düşen denek sayısının 50'den büyük olduğu durumlarda kullanılabileceğini belirtmişlerdir.

Çalışma kapsamında tanıtılan, Brunner ve arkadaşları tarafından geliştirilen yöntem, uzunlamasına verilerin analizinde ortaya çıkabilecek farklı deneme tasarımlarında, herhangi bir test varsayımı ve örneklem büyüklüğü kısıtlaması olmadan kullanım avantajına sahiptir. Bu yöntem kullanılarak elde edilen sonuçların göreceli etkiler yardımıyla yorumlanması, özellikle sağlık alanında yapılan faktöriyel çalışmalarda, klinik değerlendirmelerin yapılması ve sonuçların yorumlanması bakımından araştırmacılara kolaylık sağlayacağı düşünülmektedir.

5. KAYNAKLAR

Bonferroni, Abdi, H. , Šidák, 2007. Corrections for Multiple Comparisons. In: Salkind NJ ed. Encyclopedia of Measurement and Statistics. 1st ed. Thousand Oaks, CA: Sage; p.103-107.

Akritas, M., G., Brunner, E., 1997. A Unified Approach to Ranks Tests in Mixed Models. Journal of Statistical Planning and Inference. 61 (2): 249 - 77.

Albus, C., Theissen, P., Hellmich, M., Griebenow, R., Wilhelm, B., Aslim, D., et al., 2009. Long term Effects of a Multimodal Behavioral Intervention on Myocardial Perfusion – a Randomized Controlled Trial. International Journal of Behavioral Medicine. 16 (3): 219 - 26. DOI 10.1007/s12529-008-9030-4.

Ateş, C., 2010. Nonparametric Analysis of Longitudinal Data in Factorial Experiments. M.Sc. Thesis. Ankara Üniversitesi Sağlık Bilimleri Enstitüsü, p. 66-70.

Brunner, E., 2002. Domhof S, Langer F. Nonparametric Analysis of Longitudinal Data in Factorial Experiments. 1st ed. New York: John Wiley & Sons; p. 21-32, 35-42, 67-80.

Brunner, E., Neumann, N., 1982. Rank Tests for Correlated Random Variables. Biometrical Journal, 24 (4): 373 - 89.

Brunner, E., Puri, M., L., 2002. A Class of Rank-score Tests in Factorial Designs. *Journal of Statistical Planning and Inference*. 103 (1-2): 331- 60.

Brunner, E., Puri M., L., 2001 Nonparametric Methods in Factorial Designs. *Statistical papers*, 42 (1): 1-52.

Duruöz, M., T., Poiraudau, S., Fermanian, J., et al., 1996. Development and Validation of a Rheumatoid Hand Functional Disability Scale That Assesses Functional Handicap. *Journal of Rheumatology*, 23 (7): 1167 - 72.

Gao, O., H., Holmen, B., A., Niemeier, D., A., 2005. Nonparametric Factorial Analysis of Daily Weigh-in-motion Traffic: Implications Fort the Ozone “Weekend Effect” in Southern California. *Atmospheric Environment*, 39 (9): 1669 - 82.

Gao, Y., 2004. Rank Based Methods in the Factorial Designs for Longitudinal Data with Application to Liver Cancer Studies. M.sc. Thesis. Graduate Faculty of Texas Tech University, p. 36 - 45.

Hafız, M., 2010. Impact of Corticosteroid Phonophoresia in Carpal Tunnel Syndrome on Pain and Hand Functions: Randomised Controlled Study. Phd Thesis. Ankara Üniversitesi Tıp Fakültesi Fiziksel Tıp ve Rehabilitasyon Anabilim Dalı, p. 10 - 25.

Hollander, M., Pledger, G., Lin, P., E., 1974. Robustness of the Wilcoxon Test to a Certain Dependency Between Samples. *Ann. Statist.* 2 (1):177- 81.

Kamburoglu, K., Murat, S., Yüksel, S. P., Cebeci, A. R., Paksoy, C. S., 2010. Occlusal Caries Detection by Using a Cone-beam CT with Different Voxel Resolutions and a Digital Intraoral Sensor. *Oral Surg. Oral Med. Oral Pathol, Oral Radiol Endod.*, 109 (5) : 63-9. DOI: 10.1016/j.tripleo.2009.12.048.

Kruskal, W., H., 1952. A Nonparametric Test for the Several Sample Problems. *Annals of Mathematical Statistics*, 23 (4): 525 - 40.

Munzel, U., Bandelow, B., 1998. The Use of Parametric vs. Nonparametric Tests in the Statistical Evaluation of Rating Scales. *Pharmacopsychiatry*, 31 (6): 222 - 24.

Munzel, U., Brunner, E., 2000. Nonparametric Methods in Multivariate Factorial Designs. *Journal of statistical Planning and Inference*, 88 (1): 117 - 32.

Ruymgaart, F., H., 1980. A Unified Approach to the Asymptotic Distribution Theory of Certain Midrank Statistics. In J.P. Raoult ed. *Statistique non Parametrique Asymptotique*. Lecture Notes on Mathematics No: 821. Berlin, Springer; p. 1-18.

Shah, D., A., Madden, L., V., 2003. Nonparametric Analysis of Ordinal Data in Design Factorial Experiments. *Phytopathology*, 94 (1): 33-43.

Thompson, G., L., 1991. A Unified Approach to Rank Tests of Multivariate and Repeated Measures Designs. *Journal of the American Statistical Association*, 86 (414): 410 9.

Thompson, G., L., 1990. Asymptotic Distribution of Rank Statistics Under Dependencies with Multivariate Application. *Journal of Multivariate Analysis*, 33 (2): 183-211.

NONPARAMETRIC APPROACH FOR FACTORIAL DESIGNS WITH LONGITUDINAL DATA

ABSTRACT

Although parametric methods are commonly used for longitudinal data, the use of these methods depends on several assumptions. In recent years, non-parametric methods have been developed to investigate the effect of multiple factors for the longitudinal data that do not meet parametric test assumptions and/or that include ordinal variables. The aim of this study is to introduce the nonparametric method for the analysis of longitudinal data which was developed by Brunner and colleagues and to implement it on a data set from the field of physical medicine and rehabilitation. Marginal distributions are used for obtaining parameter estimates and hypothesis testing, in the analysis of longitudinal data with nonparametric methods. While ANOVA or Wald type statistics are used to test the hypotheses depending on the sample size, relative treatment/marginal effects are used as descriptive statistics. For the application, Duruöz Hand Index (DHI) scores are compared between two independent groups (control and treatment) at three time points; pre-treatment, during treatment and after 3 months of treatment, by using the FI-LD-FI design. The method developed by Brunner and colleagues offers solutions to different trial designs which may arise in the analyses of longitudinal data. Not requiring any test assumptions and any constraints in terms of sample size can be summarized as the advantages of this method.

Keywords: Nonparametric methods, Repeated measures, Longitudinal studies.