



ÇAPRAZ TASARIMIN KLİNİK ARAŞTIRMALARDA UYGULANMASI (APPLICATION OF CROSSOVER DESIGN IN CLINICAL RESEARCHES)

Özgür ARMANERİ*, Sergül ARMANERİ**

ÖZET/ABSTRACT

Çapraz tasarım, her bir deney birimine belirli bir sıra ile uygulanan denemeleri içeren deneyleri açıklamaktadır. Bu tasarımda bir deneme, belirli bir zaman periyodunda bir deney birimine uygulandıktan sonra aynı birime bir sonraki periyotta başka bir deneme uygulanmaktadır. Denemeler deneyin uygulandığı birim, tüm denemeleri alıncaya kadar ardışık olarak uygulanır. En yaygın çapraz tasarım türü 2-sıra, 2-periyot, 2-deneme çapraz tasarımdır. Çapraz tasarımlar, çok sayıda klinik ve farmakolojik denemelerde seçilen en yaygın tasarımlar olmuştur. Bu makalenin amacı çapraz tasarımlar hakkındaki temel bilgileri gözden geçirmek, çapraz tasarımların klinik araştırmalarda nasıl uygulandığını açıklamak, ve bir klinik denemede gerçekleştirilen çapraz tasarım uygulaması ve sonuçları hakkında bilgi vermektir.

The crossover design describes experiments with treatments administered in sequence to each experimental unit. A treatment is administered to an experimental unit for a specific period of time after which another treatment is administered to the same unit. The treatments are successively administered to the unit until it has received all treatments. The most popular crossover design is the 2-sequence, 2-period, 2-treatment crossover design. Crossover designs have been the most popular designs of choice in many clinical and pharmaceutical trials. The aim of this paper is to scrutinize the basic information about crossover designs, to explain how to use crossover designs in clinical researches, and to inform about crossover design application and its results realized on one of the clinical trial.

ANAHTAR KELİMELELER/KEYWORDS

Çapraz tasarım, Klinik denemeler, Deney tasarımı
Crossover design, Clinical trials, Design experiment

* Dokuz Eylül Üniversitesi, Mühendislik Fakültesi, Endüstri Müh. Bölümü, Bornova, İZMİR

** Aksigorta Ege Bölge Müdürlüğü, Fevzi Paşa Bulvarı, No:172/1-A, Basmane, İZMİR

1. GİRİŞ

Deney tasarımı yöntemleri süreç geliştirmede ve süreçte aksayan noktaların belirlenmesinde ve süreç performansının artırılmasında önemli rol oynarlar. Gerçekte deney düzenleme ile birtakım bilimsel süreçlerin ve sistemlerin nasıl işlediği belirlenmiş olur. Deney tasarımının üç temel prensibi olan yinleme, rasgelelik ve bloklama her deneyin en önemli bölümünü oluşturmaktadır (Montgomery, 1991).

Çapraz tasarımlar iki bloklu Latin kare tipi tasarımlardan biridir. Bu tasarımların da bir deney tasarımı olmasından dolayı çapraz tasarımların kullanıldığı deneylerde de yukarıda belirtilen temel deney tasarım ilkelerinin aynı şekilde uygulanması, elde edilen sonuçların doğruluğunu ve tutarlılığını arttıracaktır (Hinkelmann ve Kempthorne, 1994).

Günümüzde çapraz tasarımlar, klinik araştırmalar başta olmak üzere tarım ile zooloji alanlarında, otomotiv ve lastik endüstrisinde ve çeşitli alternatiflerin karşılaştırılmasını içeren diğer endüstri alanlarında kullanılmaktadır. Uygulama alanlarının giderek artması sebebiyle, bu çalışmanın çapraz tasarım literatürüne yararlı katkılar yapması amaçlanmaktadır. Çalışma şu şekilde düzenlenmiştir. İkinci bölümde çapraz tasarımlar hakkında temel bilgiler gözden geçirildikten sonra üçüncü bölümde bu tasarımların analizi üzerinde durulmuştur. Dördüncü bölümde klinik araştırmalarda çapraz tasarımların önemi konusu ele alınmış ve beşinci bölümde bir klinik araştırma probleminde çapraz tasarımın nasıl uygulandığı ve sağladığı avantajlar incelenmiştir. Son bölümde çalışmadan elde edilen sonuçlar aktarılmıştır.

2. ÇAPRAZ TASARIMLAR

Çapraz tasarım, belirli sayıdaki deney biriminin her birine iki veya daha fazla denemenin belirli bir sıra ile uygulandığı bir deney tasarımı türüdür. Çapraz tasarımlar ilk olarak tarım, çiftçilik, biyolojik tahliller, klinik denemeler, psikolojik deneyler ve hava değişimi deneyleri gibi çok farklı alanlarda kullanılmıştır (Cochran, 1939; Cochran vd., 1941; Finney, 1965; Grizzle, 1965; Keppel, 1973; Mielke, 1974; Jones ve Kenward, 1989). Çapraz tasarım denemelerinde bir deneme, belirli bir zaman diliminde deney birimine uygulanır. İzleyen periyotta aynı deney birimine bir başka deneme uygulanır. Denemelerin deney birimlerine uygulanışı, her birime tüm denemeler uygulanıncaya kadar devam eder (Kuehl, 2000).

Denemeler aynı deney birimine uygulandığı zaman, deney birimlerinin farklılığından dolayı oluşabilecek deneysel hatalar ortadan kaldırılmış olur. Böylelikle çapraz tasarımın başlıca avantajı, denemelerin karşılaştırılmalarının daha doğru, kesin ve tutarlı bir şekilde yapılmasını sağlamasıdır.

Brown, çapraz tasarımlar ile diğer tasarımlar hakkında sonuçların kesinliği, tutarlılığı ve maliyetler açısından sağlanacak tasarruflar ile ilgili araştırmalar yapmıştır (Brown, 1980). Bu araştırmalar sonucunda çapraz tasarımların çoğu klinik denemede kullanılmasının önemli avantajlar sağlayabileceğini, ancak bu tasarımların uygulanabilmesi için denemelerin etkilerinin izleyen periyotta devam etmemesi koşulunun sağlanması gerektiğini tespit etmiştir.

Çapraz tasarımlar, eğer çalışma için sınırlı sayıda deney birimi var ise bu durumda ekonomik kaynak kullanımı sağlamaktadır. Çünkü bu tasarımlar diğer tasarımlara göre aynı sayıdaki gözlem değerinin daha az deney birimi kullanarak elde edilmesini sağlamaktadır.

Son yıllarda çapraz tasarımlar, klinik denemelerde kronik rahatsızlıkların incelenmesinde sıklıkla kullanılmaktadır (Carriere ve Reinsel, 1992; Grender ve Johnson, 1993; Carriere, 1994; Donev, 1998; Carriere ve Huang, 2000). Çünkü çapraz tasarımlar, paralel grup tasarımlara göre daha az deney birimi (hasta) gerektirir. Paralel grup tasarımlarda her deney birimine sadece bir deneme uygulanır. Ayrıca gözlem değerleri üzerindeki, bir deney

biriminden diğerine geçişten kaynaklanan değişkenlik çapraz tasarım ile yok edildiğinden denemelerin doğrudan etkilerinin tahminleri çok daha etkin şekilde belirlenebilir.

2.1. Tasarım Özellikleri

Bir çapraz tasarım denemesinde, denemelerin uygulama sıralaması *sıra*, bir denemenin uygulama zamanı ise *periyot* olarak adlandırılır. Denemeler genellikle A, B gibi büyük harflerle gösterilir. Tasarımda sıralar, deney gerçekleştirilmeden önce belirlenmiş olmalı ve deney birimleri bu sıralara *rasgele* atanmalıdır.

En yaygın kullanılan çapraz tasarım türü, 2x2 çapraz tasarım olarak da adlandırılan 2-sıra, 2-periyot, 2-deneme çapraz tasarımıdır. Bu tür tasarımların dışında, farklı amaçlara yönelik denemelerde kullanılan ve daha fazla sayıda deneme içeren çapraz tasarım türleri mevcuttur (Senn, 1993; Kuehl, 2000).

2.2. Aktarılmış Etkiler

Çapraz tasarımın en önemli dezavantajlarından biri, bir periyotta uygulanan denemenin deney birimi üzerindeki etkisinin, takip eden periyotta da devam etmesidir. Bir denemenin, uygulandığı periyodu takip eden periyotta devam eden etkisine *aktarılmış etki* denir. Bir periyotta elde edilen gözlem değeri üzerinde aktarılmış etkinin varlığı, o periyotta uygulanan denemenin doğrudan etkisinin belirlenmesini imkansız hale getirir. Bu nedenle olası aktarılmış etkiler ihmal edilmemeli ve mutlaka analiz edilmelidir (Bose, 2002).

Çapraz tasarım uygulamalarında, denemelerin deney birimleri üzerindeki etki sürelerinin belirlenmesi gerekir. Bir deney biriminin, kendine uygulanan denemenin etkisinden kurtularak deneme uygulanmadan önceki durumuna geri dönmesi için geçen süreye *dinlenme süresi* adı verilir. Çapraz tasarımda, ancak bu süre geçtikten sonra deney birimine farklı bir deneme uygulanır. Dinlenme süresinin uzunluğu, aktarılmış etkilerin varlığını da yakından ilgilendirir.

j .periyotta elde edilen bir gözlem değeri üzerinde 1,..., $j-1$. periyotlarda uygulanan denemelerin aktarılmış etkilerinin var olması mümkündür. Ancak genellikle çapraz tasarımlarda, herhangi bir periyottaki gözlem değeri üzerinde, sadece bir önceki periyottan gelebilecek aktarılmış etkinin varlığı araştırılmaktadır (Chinchilli, 2003).

Çapraz tasarımlar ile ilgili daha ayrıntılı bilgi için literatürde yer alan çalışmalar incelenebilir (Jones ve Kenward, 1989; Senn, 1993; Kuehl, 2000; Chinchilli, 2003).

3. ÇAPRAZ TASARIMLARIN ANALİZİ

3.1. 2-Deneme, 2-Periyot Çapraz Tasarımların Analizi

Aşağıdaki analiz sürecinde yanıt değişkeninin normal dağılım gösteren sürekli rasgele değişken olduğu varsayılır. Süreçte denemeler, R ve T harfleri ile gösterilecektir. Süreç aşağıdaki şekilde özetlenmiştir (Chow ve Liu, 1999). Buna göre 2x2 çapraz tasarım için öngörülen genel doğrusal model aşağıdaki gibi gösterilebilir.

$$Y_{ijk} = \mathbf{m} + S_{ik} + P_j + F_{(j,k)} + C_{(j-1,k)} + e_{ijk} \quad (1)$$

Modelde; i , deney birimini ($1'$ den n_k 'ya kadar); j , uygulama periyodunu (1 ya da 2) ve k , sıra numarasını (1 ya da 2) göstermektedir. Ayrıca, \mathbf{m} genel ortalamayı; S_{ik} , k .sıradaki i .deney biriminin rasgele etkisini; P_j , j .periyodun gözlem değeri üzerindeki etkisini; $F_{(j,k)}$, k .sıranın

j .periyodunda uygulanan denemenin doğrudan etkisini, ve $C_{(j-1,k)}$, k .sırada $(j-1)$.periyotta uygulanan bir denemenin j .periyotta devam eden aktarılmış etkisini e_{ijk} ise rasgele hata bileşenini göstermektedir. Sonuç olarak Y_{ijk} ise k .sırada j .periyotta i .deney birimine ait gözlem değerini gösterir. Deney birimleri sıralara rasgele atandıklarından, k .sıradaki i .deney biriminin rasgele etkisinin (S_{ik}) ortalamasının sıfır olduğu varsayılır.

Bu bilgiler ışığında, 2x2 çapraz tasarımın şu şekilde gösterilmesi mümkündür (Çizelge 1).

Çizelge 1. 2x2 çapraz tasarımda beklenen ortalama değerler

Sıra	Periyot 1	Periyot 2
1 (RT)	$m_{11} = m + P_1 + F_R$	$m_{21} = m + P_2 + F_T + C_R$
2 (TR)	$m_{12} = m + P_1 + F_T$	$m_{22} = m + P_2 + F_R + C_T$

2x2 çapraz tasarımlar aktarılmış etkilerin olmadığı durumlarda kullanıldığından ilk olarak bu etkilerin varlığı analiz edilmelidir. Analiz sonucunda aktarılmış etkilerin varlığı ortaya çıkarsa, tasarım yeniden gözden geçirilmeli, dinlenme sürelerinin değiştirilerek en uygun dinlenme süresinin belirlenmesine çalışılmalıdır. Analiz sonucunda aktarılmış etkilerin olmadığı sonucuna ulaşılması durumunda diğer etkilerin varlığının analizine geçilebilir.

Aktarılmış etkilerin varlığı, $C_T = C_R = 0$ olup olmadığı hipotezinin, t testi kullanılarak test edilmesi ile belirlenebilir. Bu durumda test istatistiği

$$T_c = \frac{\hat{C}}{\hat{S}_u \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (2)$$

olacaktır. Bu istatistikte \hat{C} , denemelerin aktarılmış etkileri arasındaki farkın ($C_T - C_R$) tahminleyicisidir. Paydada ise bu tahminleyicinin standart hata değeri görülmektedir. n_k değerleri, sıralardaki deney birimi sayısını göstermektedir. Buna göre T_c test istatistiği değeri aşağıdaki formüller kullanılarak hesaplanmaktadır.

$$\hat{C} = \bar{U}_{.2} - \bar{U}_{.1} \quad (3)$$

$$\bar{U}_{.k} = \frac{1}{n_k} \sum_{i=1}^{n_k} U_{ik} \quad (4)$$

$$\hat{S}_u^2 = \frac{1}{(n_1 + n_2 - 2)} \sum_{k=1}^2 \sum_{i=1}^{n_k} (U_{ik} - \bar{U}_{.k})^2 \quad (5)$$

$$U_{ik} = Y_{i1k} + Y_{i2k} \quad (6)$$

$H_0: C_T = C_R = 0$ hipotezi, α önem düzeyinde $|T_c| > t_{\alpha/2, n_1+n_2-2}$ koşulunun sağlanması durumunda reddedilir. $C = C_T - C_R$ için % 100(1- α) güven aralığı aşağıdaki gibidir.

$$\hat{C} \pm (t_{\alpha/2, n_1+n_2-2}) \hat{S}_u \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad (7)$$

Denemelerin doğrudan etkilerinin varlığı, $F_T = F_R = 0$ olup olmadığı hipotezinin, t testi kullanılarak test edilmesi ile belirlenebilir. Bu durumda test istatistiği

$$T_d = \frac{\hat{F}}{\hat{S}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (8)$$

İstatistikte \hat{F} , denemelerin doğrudan etkileri arasındaki farkın ($F_T - F_R$) tahminleyicisidir. Buna göre T_d test istatistiği değeri aşağıdaki formüller kullanılarak hesaplanmaktadır.

$$\hat{F} = \bar{d}_{.1} - \bar{d}_{.2} \quad (9)$$

$$\bar{d}_{.k} = \frac{1}{n_k} \sum_{i=1}^{n_k} d_{ik} \quad (10)$$

$$\hat{S}_d^2 = \frac{1}{(n_1 + n_2 - 2)} \sum_{k=1}^2 \sum_{i=1}^{n_k} (d_{ik} - \bar{d}_{.k})^2 \quad (11)$$

$$d_{ik} = \frac{Y_{i2k} - Y_{i1k}}{2} \quad (12)$$

$H_0: F_T = F_R = 0$ hipotezi, α önem düzeyinde $|T_d| > t_{\alpha/2, n_1+n_2-2}$ koşulunun sağlanması durumunda reddedilir. $F = F_T - F_R$ için % 100(1- α) güven aralığı aşağıdaki gibidir.

$$\hat{F} \pm (t_{\alpha/2, n_1+n_2-2}) \hat{S}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad (13)$$

Periyot etkisinin varlığı, $P_1 = P_2 = 0$ olup olmadığı hipotezinin, t testi kullanılarak test edilmesi ile belirlenebilir. Bu durumda test istatistiği

$$T_p = \frac{\hat{P}}{\hat{S}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (14)$$

Bu istatistikte \hat{P} , periyot etkileri arasındaki farkın ($P_2 - P_1$) tahminleyicisidir. Buna göre T_p test istatistiği değeri aşağıdaki formüller kullanılarak hesaplanmaktadır.

$$\hat{P} = \bar{O}_{.1} - \bar{O}_{.2} \quad (15)$$

$$\bar{O}_{.1} = \bar{d}_{.1} \quad (16)$$

$$\bar{O}_{.2} = -\bar{d}_{.2}$$

$$\hat{S}_d^2 = \frac{1}{(n_1 + n_2 - 2)} \sum_{k=1}^2 \sum_{i=1}^{n_k} (d_{ik} - \bar{d}_{.k})^2 \quad (17)$$

$$d_{ik} = \frac{Y_{i2k} - Y_{i1k}}{2} \quad (18)$$

$H_0: P_1=P_2=0$ hipotezi, α önem düzeyinde $|T_p| > t_{\alpha/2, n_1+n_2-2}$ koşulunun sağlanması durumunda reddedilir. $P=P_2-P_1$ için % 100(1- α) güven aralığı aşağıdaki gibidir.

$$\hat{P} \pm (t_{\alpha/2, n_1+n_2-2}) \hat{S}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad (19)$$

Denemelerin doğrudan etkilerinin bulunup bulunmadığını test ederken kullanılan t testi, bu denemelerin biyolojik olarak eşit etkiye sahip olup olmadığını bir ön değerlendirmesi olarak düşünülebilir. Kullanılan t testi, iki denemenin etki yönünden birbirinden farklı olup olmadığını araştırmaktadır. Bu test, iki tedavinin biyolojik olarak eşdeğerliğini değerlendirmez. Denemelerin ortalama doğrudan etkisinin eşitliği hipotezini reddedememek, denemelerin biyolojik olarak etkilerinin birbirinin aynı olduğu anlamına gelmez.

% ± 10 kuralı biyolojik eşitlik testlerinde kullanılmaktadır. Bu kural, iki denemeye ilişkin ortalama yanıt değerleri arasındaki farkın, referans olarak alınan denemenin yol açtığı ortalama yanıt değerinin % ± 10 'u olmasına izin vermektedir. % ± 10 kuralı, iki denemeye ilişkin ortalama yanıt değerlerinin oranı (m_T/m_R) için de kullanılabilir. Bu durumda biyolojik eşitlik sınırları 0.9 ve 1.1 aralığı olacaktır. Yani eğer m_T/m_R oranı için bulunan güven aralığı, 0.9 ile 1.1 aralığı içine düşüyorsa, bu durumda iki denemenin biyolojik olarak birbirinin eşdeğeri olduğunu söylemek mümkündür. Biyolojik eşitliğin testi için çok çeşitli yöntemler önerilmiştir. Bunların başında *Güven Aralığı Yaklaşımı* gelmektedir (Weslake, 1981). Bu yaklaşım, $m_T - m_R$ farkı ya da m_T/m_R oranı için bulunan % $(1-2\alpha) \times 100$ güven aralığının biyolojik eşitlik kabul limitleri içinde olması durumunda biyolojik eşitliğin sağlanacağını öngörür. % ± 10 kuralı uygulandığında, $m_T - m_R$ farkı için bulunan güven aralığının % $\pm m_R \times 10$ aralığında olması gerekmektedir. Aynı şekilde m_T/m_R oranı için bulunan güven aralığı da 0.9 ile 1.1 aralığı içinde yer almalıdır.

3.2. t -Deneme, p -Periyot Çapraz Tasarımların Analizi

Bu tip çapraz tasarımlarda n tane sıra ve k sırada da n_k tane deney birimi bulunmaktadır. t tane deneme olup, her bir deney birimi grubuna p periyot boyunca bu denemeler farklı sıralarda uygulanmaktadır. Y_{ijk} ; k sırada, j periyotta i deney birimine ait gözlem değerini ifade eder. Bu durumda doğrusal model aşağıdaki gibi gösterilir.

$$Y_{ijk} = \mu + T_k + S_{ik} + P_j + F_{(j,k)} + C_{(j-1,k)} + e_{ijk} \quad (20)$$

$$i=1,2,\dots,n \quad j=1,2,\dots,p \quad k=1,2,\dots,n$$

Bu modeldeki notasyonlar, 2×2 çapraz tasarım için olan doğrusal modelde kullanılanlar ile aynıdır. Tek farklılık, modele T_k teriminin eklenmesidir. Bu terim, deney birimine k sıranın uygulanmasının gözlem değeri üzerindeki etkisini gösterir.

t -deneme, p -periyot çapraz tasarımları analiz etmek için varyans analizi kullanılabilir. Bu çapraz tasarım modeli için varyans analizi tablosu Çizelge 2'de görülmektedir (Kuehl, 2000).

Çok sayıda istatistiksel bilgisayar yazılımı, ihtiyaç duyulan kareler toplamı değerlerini hesaplar. Bu değerlerin el ile hesaplanmasında kullanılan formüller geliştirilmiştir (Petersen, 1985; Kuehl, 2000).

Çizelge 2. *t*-deneme, *p*-periyot çapraz tasarımlar için varyans analizi $N = \sum_k^n n_k$ (Kuehl, 2000)

Değişim Kaynağı	Serbestlik Derecesi	Kareler Toplamı	Kareler Ortalaması	F
Toplam	$Np-1$	SSToplam	$MSToplam = SSToplam/(Np-1)$	
Denemeler Arası:				
Sıra	$n-1$	SSS	$MSS = SSS/(n-1)$	
Sıradaki deney birimleri	$(N-n)$	SSW	$MSW = SSW/(N-n)$	
Denemeler İçi:				
Periyot	$p-1$	SSP	$MSP = SSP/(p-1)$	$F_P = MSP/MSE$
Denemeler (doğrudan)	$t-1$	SST	$MST = SST/(t-1)$	$F_F = MST/MSE$
Denemeler (aktarılmış)	$t-1$	SSC	$MSC = SSC/(t-1)$	$F_C = MSC/MSE$
Hata	$(N-1)(p-1)-2(t-1)$	SSE	$MSE = SSE/[(N-1)(p-1)-2(t-1)]$	

Bu tip tasarımlarda, aktarılmış etkilerin, denemelerin doğrudan etkilerinin ve periyot etkilerinin varlığı, Çizelge 3'te görülen hipotez testleri kullanılarak belirlenebilir.

Çizelge 3. *t*-deneme, *p*-periyot çapraz tasarımlar için hipotez testleri

Etki Türü	H ₀ Hipotezi	Test İstatistiği	Koşul	Karar
Aktarılmış Etkiler	Tüm $C_{(j-1,k)} = 0$	$F_C = MSC/MSE$	$F_C > F_{\alpha, (t-1), [(N-1)(p-1)-2(t-1)]}$	H ₀ red
Denemelerin Doğrudan Etkisi	Tüm $F_{(j,k)} = 0$	$F_F = MST/MSE$	$F_F > F_{\alpha, (t-1), [(N-1)(p-1)-2(t-1)]}$	H ₀ red
Periyot Etkisi	Tüm $P_j = 0$	$F_P = MSP/MSE$	$F_P > F_{\alpha, (p-1), [(N-1)(p-1)-2(t-1)]}$	H ₀ red

4. KLİNİK ARAŞTIRMALARDA ÇAPRAZ TASARIMLAR

Çapraz tasarımlar, çoğu klinik ve farmakolojik denemede, çok çeşitli denemelerin birbirleri ile karşılaştırılmasında seçilen en yaygın tasarımlar olmuştur. Bu tasarımların araştırmacılar tarafından tercih edilmelerinin çeşitli sebepleri bulunmaktadır. Bu nedenlerin başlıcaları; denemeye yeni hastaların eklenmesindeki bütçe kısıtlamaları, deneye katılacak hastaların bulunmasında karşılaşılan güçlükler ve denemeye katılan her yeni hastanın, denemeye ilgili özel eğitimlerindeki zaman kısıtlarıdır (Carriere ve Huang, 2000).

Çapraz tasarımlarda denemelerin belirli sıra ile aynı hastaya uygulanması, bu denemelerin farklı farklı hastalara uygulanması sonucu gözlem değerinde ortaya çıkacak olan ve hastaların farklı fiziksel veya psikolojik özelliklerinden kaynaklanan değişkenliği yok etmektedir. Bu da, denemelerin karşılaştırılmasında doğru sonuçlara ulaşma başarısını artırır.

Özetle çapraz tasarım denemelerinin, paralel grup denemelerine göre tercih edilmesinin iki ana sebebi bulunmaktadır (Senn, 1993):

- Aynı sayıda gözlem değeri elde edebilmek için daha az sayıda hasta yeterlidir.
- Analiz ve tahminlerde, daha az gözlem değeri ile aynı doğruluk ve tutarlılık sağlanır.

Çapraz tasarım denemeler, kronik hastalıklar için geliştirilen denemelerin karşılaştırılması için uygundur. Böyle durumlarda, denemelerin hastalıkları tamamen iyileştirici değil,

hastalığın şiddetini yatıştırıcı etkisi vardır. Çapraz tasarımların uygulanması için en uygun alanlar astım, migren, hafif ve orta derecede hipertansiyon ve sara gibi kronik hastalıklardır.

Tasarımda etkisi çok uzun süreli denemeler kullanılıyorsa, bu etkinin geçmesi için gereken dinlenme süresi de çok uzun olacaktır. Bu süre yeterince uzun verilmezse bir sonraki periyotta aktarılmış etkiler görülecektir. Bu nedenle, bu tip denemelerin karşılaştırılmasında çapraz tasarım kullanılması doğru olmaz. Yani çapraz tasarımlar, etki süresi nispeten kısa olan denemelerin karşılaştırılmasında kullanılmaktadır.

Çapraz denemeler ile paralel grup denemeler karşılaştırıldığında, çapraz denemelerin avantajları olduğu kadar dezavantajları da vardır. Bunlar şöyle sıralanabilir (Senn, 1993);

- (1) Çapraz tasarımlarda denemeden ayrılma problemleri görülebilir. Hastaların denemeden ayrılması, paralel grup denemelerin analizinde de güçlükler yol açar. Ancak çapraz tasarım denemelerinde hastanın denemeden ayrılmasının yol açtığı sorunlar daha büyük olacaktır.
- (2) Deneme periyotları boyunca denemelerin uygulandığı hastanın ölüm ihtimali söz konusu olursa veya denemeler kendine sırayla uygulandığında hasta acı çekiyorsa bu tip hastalar için çapraz tasarımlar kesinlikle uygun değildir.
- (3) Çapraz denemelerde hastaların her birine bütün denemeler uygulandığı ve aralarda da dinlenme süreleri bırakıldığı için hastalar kendilerine uygulanacak denemelerin sayılarının çok olmasından ve gözlem altında geçirecekleri sürenin uzun olmasından rahatsızlık duyabilirler.
- (4) Son olarak çapraz tasarımların analizi daha zordur. Çünkü, denemelerin doğrudan etkilerinin yanında, aktarılmış etkilerin ve periyot etkilerinin varlığı da araştırılmalıdır.

Herhangi bir klinik denemenin planlama aşamasında, deneyi gerçekleştirecek kişinin karşılaşacağı temel sorun, denemede ne kadar hastaya ihtiyaç duyulacağı sorunudur. Denemede yer alacak hastaların sayısı denemenin büyüklüğü olarak adlandırılır. Bir klinik denemede amaca göre ihtiyaç duyulan hasta sayısı $n \cong [(z\alpha/2 + z\beta)2\sigma^2]/\delta^2$ eşitliği ile belirlenebilir. Bu eşitlikte; n, toplam hasta sayısını; σ^2 , tek bir hasta üzerinden elde edilen gözlem değerleri arasındaki farkların varyansını; δ , denemelerin ortalama etkileri arasındaki hedef farkın değerini; α ile β ise I. ve II. tip hata olasılıklarını göstermektedir. Deneme büyüklüğünün belirlenmesinde kullanılan parametreler, deneyin amaçları ve bu alanda daha önce yapılan araştırmalar dikkate alınarak deneyi düzenleyen tarafından belirlenmektedir (Pocock, 1983).

5. BİR KLİNİK ARAŞTIRMANIN SONUÇLARININ ÇAPRAZ TASARIMLA DEĞERLENDİRİLMESİ

Bu bölümde, ilk olarak gerçek bir klinik araştırma problemi tanıtılmıştır. Daha sonra deney sonucunda elde edilen veriler, çapraz tasarım ile analiz edilmiş ve elde edilen sonuçlar hakkında bilgi verilmiştir. Çalışmada, orta ve şiddetli astım hastalığına yakalanan çocukların akciğer fonksiyonları üzerinde 2 farklı denemenin (ilacın) etkilerinin karşılaştırılması amaçlanmaktadır. Bu denemeler, daha önceden geliştirilmiş ve halen kullanılmakta olan tek doz 200 µg salbutamol ile yeni geliştirilmiş tek doz 12 µg formoterol ilaçlarıdır.

Bir sonraki adımda, iki farklı denemenin akciğer fonksiyonu üzerindeki etkisini belirlemek amacıyla bir yanıt değişkeni seçilecektir. Denemelerin yanıt değişkeni üzerindeki etkileri belirlenmek istendiğinden, tasarımda faktörler açıklanan iki farklı denemedir.

Yanıt değişkeni, inceleme altında olan çalışma hakkında gerçekten faydalı ve kullanışlı bilgiler sağlamalıdır. Yapılan tıbbi inceleme ve araştırmalar sonucunda *Maksimum Soluk Verme Akışı (litre/dakika)*'nın akciğer fonksiyonunun bir ölçüsü olduğu ve bu çalışma

hakkında gerekli bilgileri sağlayacağı belirlenmiştir. Sonuç olarak, Soluk Verme Akışı (SVA) tasarımı yanıt değişkeni olarak belirlenmiştir.

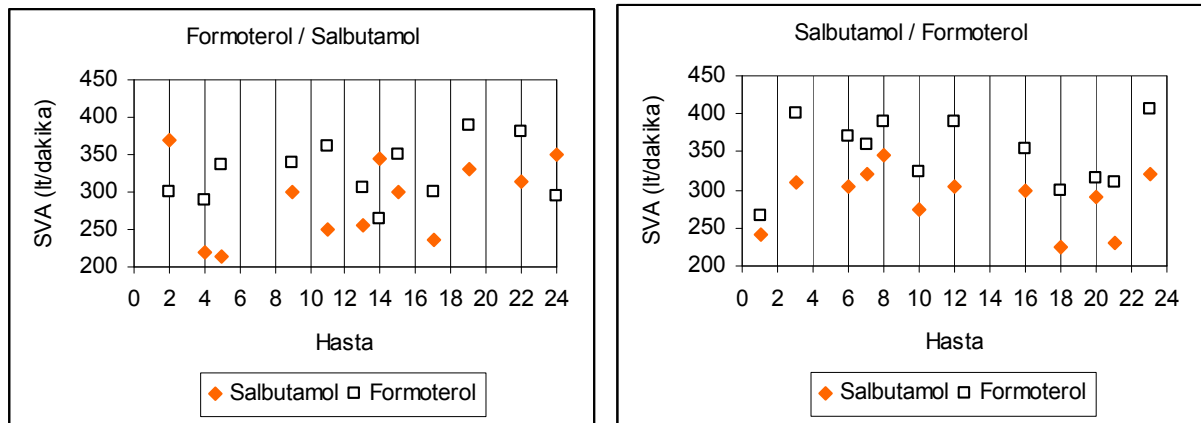
Çalışmada 2 farklı denemenin yanıt değişkeni üzerindeki etkisi belirlenmek istendiğinden tasarım türü olarak en yaygın çapraz tasarım türü olan 2×2 çapraz tasarım seçilmiştir. Tasarımda SVA değerleri, yaklaşık normal dağıldığı varsayılan sürekli raslantı değişkenidir.

5.1. Deneyin Gerçekleştirilmesi

Bu deney gerçekleştirilirken şöyle bir sistematik sürecin izlenmesi kararlaştırılmıştır; Çocuklar 2 deneme sırasından birine rasgele atanacaktır. Birinci gruba sabah 1 doz formoterol verilecek, hastalar klinikte 8 saat süresince gözlenecek ve 8 saatin sonunda SVA değeri ölçülecektir. Sonra hastalar evlerine gönderilecektir. En az 1 günlük dinlenme süresinden sonra hastalar tekrar kliniğe gelecek ve herbirine 1 doz salbutamol uygulanacak ve yine 8 saat sonunda hastaların SVA değeri ölçülecektir. Diğer grupta yer alan hastalara da kliniğe ilk geldiklerinde 1 doz salbutamol, ikinci ziyaretlerinde ise 1 doz formoterol verilecektir.

Deneyi gerçekleştirmek için öncelikle, ne kadar çocuğa ihtiyaç olduğunun belirlenmesi gerekir. Çalışmada, iki denemenin hastalar üzerinde yarattığı ortalama etki arasındaki fark eğer 30 litre/dakika ve daha üzeri olursa, bu durumda iki denemenin doğrudan etkilerinin birbirlerinden farklı olduğu kararına varılması istenmektedir. Bu durumda d değeri 30 litre/dakika olarak belirlenmiştir. Testin gücünün % 90, I. tip hata olasılığının % 5 ve geçmiş denemeler dikkate alınarak s değerinin 45 litre/dakika olduğu varsayılmıştır. Bu veriler kullanılarak denemede yer alması gereken hasta sayısı 24 çocuk olarak belirlenmiştir.

Deneyi gerçekleştirmek için tüm veriler elde edildikten sonra deney gerçekleştirilmiştir. Elde edilen SVA değerlerinin grafiksel gösterimi Şekil 1'deki gibidir.



Şekil 1. Deneyin gerçekleştirilmesi sonucu her iki sıra için elde edilen SVA değerleri

5.2. Periyot Etkisini ve Aktarılmış Etkileri İhmal Ederek Tasarımın Analizi

Bu bölümde, çalışmada olası periyot etkileri ve aktarılmış etkiler ihmal edilerek analiz yapılmıştır. Herhangi bir 2×2 çapraz tasarım denemesinde periyot etkisinin ve aktarılmış etkilerin ihmal edilmesi durumunda, denemelerin hastalara hangi sıra ile ve ne zaman uygulandığının bir önemi kalmamaktadır. Çalışmada SVA değerleri, aynı deney birimi üzerinde iki gözlem yapılarak elde edildiğinden böyle bir veriyi analiz etmek için gözlemlerin çiftler halinde elde edilmesi durumunda kullanılan eşleştirilmiş t -testi kullanılmalıdır. Eşleştirilmiş t -testi için kullanılacak hipotez; $H_0: \mu_R - \mu_T: \mu_D = 0$ ve $H_1: \mu_R - \mu_T: \mu_D \neq 0$ hipotezidir.

$\mu_D = \mu_R - \mu_T$ ifadesi, 1 doz formoterol (*R*) uygulandığında beklenen yanıt ile 1 doz salbutamol (*T*) uygulandığında beklenen yanıt arasındaki farkı belirtir. Analizin gerçekleştirilmesi için MINITAB istatistik yazılımı kullanılmıştır. Elde edilen sonuçlar aşağıda gösterilmiştir.

Paired T for C1 - C2

	N	Mean	StDev	SE Mean
C1	24	337,29	42,91	8,76
C2	24	289,58	45,75	9,34
Difference	24	47,7	50,6	10,3

95% CI for mean difference: (26,3; 69,1)

T-Test of mean difference=0 (vs not=0): T-Value=4,62 P-Value=0,000

Görüldüğü gibi, test istatistiği değeri $T=4.62$ olarak bulunmuştur. Bu değer, çizelge değerinden ($t_{0,025,23}=2.069$) büyük olduğundan yukarıdaki H_0 hipotezi reddedilir. Ortalama etkiler arasındaki fark (μ_D) için % 95 güven aralığı da aşağıdaki gibi elde edilmiştir.

$$26,3 \leq m_D \leq 69,1 \quad (21)$$

$$26,3 \leq m_R - m_T \leq 69,1$$

Dikkat edilirse, güven aralığı sıfır içermemektedir. Her iki sınır değeri de pozitif olduğundan, bu çalışmada yer alan iki denemenin hastalar üzerindeki ortalama etkileri % 95 güvenle birbirlerinden farklıdır ve 1 doz formoterol ilacının hastalar üzerinde yarattığı ortalama etki, 1 doz salbutamol'un ortalama etkisinden daha büyüktür.

5.3. Çapraz Tasarımın Analizi

Çapraz tasarım denemesinin analizi için NCSS istatistik yazılımı kullanılmıştır. Analiz için gerekli olan çapraz tasarım parametrelerinin değerleri Çizelge 4'te görülmektedir

Çizelge 4. Çapraz tasarım analizinde kullanılan parametre değerleri

Çapraz tasarım analizinde kullanılacak α düzeyi	0.05
Biyolojik eşitlik analizinde kullanılacak α düzeyi	0.05
Üst biyolojik eşitlik sınırı	10%
Alt biyolojik eşitlik sınırı	-Üst sınır

Yukarıdaki değerler, literatürde en çok kullanılan ve önerilen değerler olarak alınmıştır. Analiz prosedürünün çalıştırılması sonucunda elde edilen sonuçlar Çizelge 5'te görülmektedir. Analizlerde 1 doz *formoterol* ilacı *R* harfi ile 1 doz *salbutamol* ise *T* harfi ile gösterilmiştir.

Çizelge 5. Çapraz tasarım analizi sonuç çizelgesi

Parametre	Tahmini Etki	Standart Hata	T Değeri (SD=22)	Olasılık Düzeyi	Alt 95,0% Güven Sınırı	Üst 95,0% Güven Sınırı
$F = F_T - F_R$	-47,71	10,24	-4,66	0,0001	-68,94	-26,48
$P = P_2 - P_1$	12,29	10,24	1,20	0,2426	-8,94	33,52
$C = C_T - C_R$	21,25	30,07	0,71	0,4871	-41,11	83,61

Sonuç çizelgesine göre, incelenen 2x2 çapraz tasarım çalışmasında yer alan iki denemenin ortalama etkileri, 0,05 önem düzeyinde, birbirlerinden önemli farklılıklar göstermektedir. (gerçek önem düzeyinin 0,0001 olduğu görülmektedir). Tasarımda 12 hasta birinci sırada (RT), 12 hastada ikinci sırada (TR) yer almıştır. R denemesinin uygulanması sonucu elde edilen ortalama SVA değeri 337,29 litre/dakika iken T denemesinin uygulanması sonucu elde edilen ortalama SVA değeri 289,58 litre/dakika'dır.

Sonuç çizelgesinde, periyot etkisinin ve aktarılmış etkilerin tahmini değerleri sırasıyla 12,29 ve 21,25 litre/dakika olarak elde edilmiştir. Ancak bu parametreler için bulunmuş olan güven aralıklarının alt ve üst sınırlarına bakıldığında, her iki % 95 güven aralığı da sıfır değerini içermektedir. Ayrıca her iki parametreye ait olasılık düzeyi de 0,05 değerinden büyüktür. Bu sonuçlar göstermektedir ki; 0,05 önem düzeyinde incelenen çapraz tasarımda, periyot etkileri ve aktarılmış etkiler önemsiz düzeydedir.

Aşağıdaki çizelgede, denemenin gerçekleştirilmesi sonucunda elde edilen verilere ilişkin, analiz sürecinde kullanılan ortalama, standart sapma ve standart hata değerleri görülmektedir.

Analiz sürecinde, aktarılmış etkilerin, periyot etkisinin ve denemelerin doğrudan etkilerinin varlığı analiz edildikten sonraki aşama, denemelerin biyolojik eşitliğinin olup olmadığının test edilmesi aşamasıdır. Daha önce aktarılmış etkilerin ve periyot etkisinin gözlem değerleri üzerindeki etkilerinin önemsiz olduğu ve denemelerin doğrudan etkilerinin birbirlerinden önemli düzeyde farklı olduğu sonucuna varılmıştı. Bu sonuçlar, denemelerin biyolojik olarak eşit olmadıklarını açıkça göstermektedir. Bununla birlikte, analiz sürecinde yer alması ve bulguların doğruluğunun kontrolü açısından biyolojik eşitlik testi de yapılmış ve aşağıdaki sonuçlar elde edilmiştir.

Çizelge 6. Verilere ilişkin ortalama, standart sapma ve standart hata değerleri

Sıra	Periyot	Tedavi	Sayı	EKK Ortalaması	Standart Sapma	Standart Hata
1	1	R	12	325,83	39,07	11,28
2	2	R	12	348,75	45,13	13,03
1	2	T	12	290,42	53,79	15,53
2	1	T	12	288,75	38,50	11,11
1	Fark	(T-R)/2	12	-17,71	33,48	9,67
2	Fark	(T-R)/2	12	30,00	11,68	3,37
1	Toplam	R+T	12	616,25	65,99	19,05
2	Toplam	R+T	12	637,50	80,58	23,26
.	.	R	24	337,29	-	8,62
.	.	T	24	289,58	-	9,55
1	.	.	24	308,13	-	-
2	.	.	24	318,75	-	-
.	1	.	24	307,29	-	7,92
.	2	.	24	319,58	-	10,13

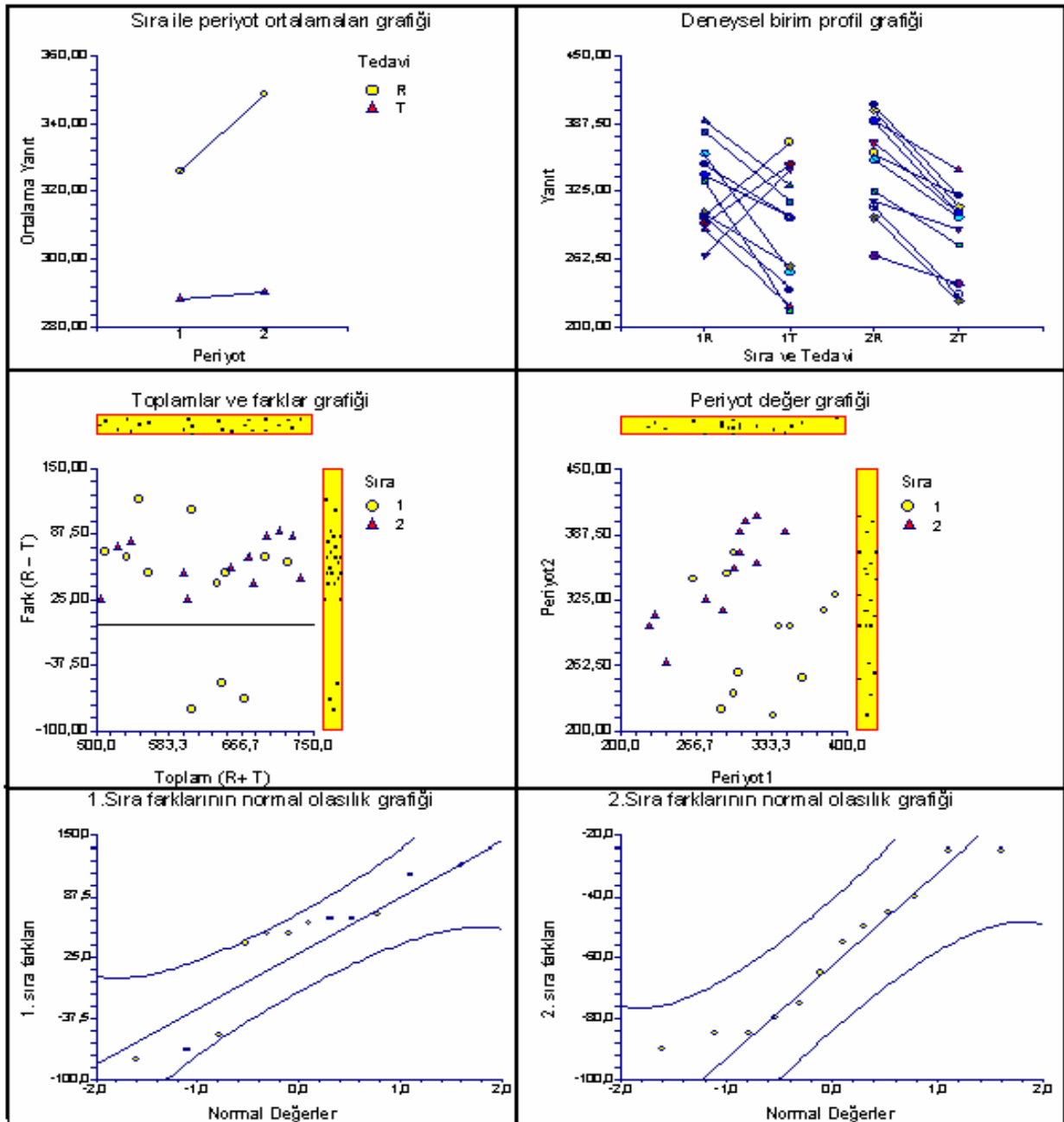
Çizelge 7. $\mu_T - \mu_R$ farkı için güven aralığını temel alan biyolojik eşitlik testi sonuçları

Test Tipi	Alt Biyolojik Eşitlik Sınırı	Alt % 90 Güven Aralığı Sınırı	Üst % 90 Güven Aralığı Sınırı	Üst Biyolojik Eşitlik Sınırı	% 5 önem düzeyinde denemeler eşit mi?
$m_T - m_R$ farkı için güven aralığı	-33,73	-65,29	-30,13	33,73	Hayır

Çizelge 8. μ_T / μ_R farkı için güven aralığını temel alan biyolojik eşitlik testi sonuçları

Test Tipi	Alt Biyolojik Eşitlik Sınırı	Alt % 90 Güven Aralığı Sınırı	Üst % 90 Güven Aralığı Sınırı	Üst Biyolojik Eşitlik Sınırı	% 5 önem düzeyinde denemeler eşit mi?
m_T / m_R oranı için güven aralığı	90,00	80,64	91,07	110,00	Hayır

Her iki sonuç çizelgesinin son sütununda da görüldüğü gibi, denemelerin ortalama etkileri % 5 önem düzeyinde birbirine eşit değildir. Biyolojik eşitlik testi sonucunda elde edilen sonuçlar daha önceden ulaşılan sonuçları desteklemekte ve doğrulamaktadır.



Şekil 2. Çapraz tasarım analizi sonuç grafikleri

Son olarak, incelenen veri grubuyla ilgili tüm değerleri şekilsel olarak belirlemeye yarayan grafikler elde edilmiştir. Analiz sonucunda elde edilen grafikler Şekil 2’de görülmektedir.

Sıra ile periyot ortalamaları grafiğinde, doğruların uzunluğu denemenin doğrudan etkisinin büyüklüğünü göstermektedir. Bu durumda R ile adlandırılan formoterol ilacının ortalama etkisinin, T ile adlandırılan salbutamol ilacından daha büyük olduğu söylenebilir. İki doğrunun kesişmemesi ise aktarılmış etkilerin ve periyot etkisinin önemsiz olduğunu göstermektedir. Deneysel birim profil grafiğinde, gözlem değerleri arasında, analiz dışı bırakılması gereken herhangi bir uç değer görülmemektedir. Toplamlar ve farklar grafiğinde, sembollerin yatay eksen boyunca ayrılma göstermeleri aktarılmış etkinin varlığını gösterir. Burada, semboller birlikte hareket ettiklerinden yine önemli bir aktarılmış etkinin olmadığını söylememiz mümkündür. Sembollerin düşey ekseninde ayrılma göstermesi denemelerin doğrudan etkilerinin büyüklüğünü göstermektedir. Grafikte, ister birinci sırada, isterse ikinci sırada olsun, $R-T$ farkı çoğunlukla pozitif çıkmıştır. Bu nedenle, yine formoterolün ortalama etkisinin salbutamole göre daha fazla olduğu sonucuna varılabilir. Periyot değer grafiğinde de hangi sırada olursa olsun, formoterol uygulanan periyotlardaki SVA değerlerinin, salbutamol uygulanan periyotlarda elde edilen değerlere göre çoğunlukla daha büyük olduğu görülmektedir.

Çapraz tasarım analizi, aynı deney birimleri üzerinden alınan gözlem değerleri arasındaki farklar yaklaşık normal dağılım gösterirse, doğru sonuçlar verecektir. Bundan dolayı, tasarımda her iki sıra için elde edilen farkların normal dağılım gösterip göstermediği de test edilmiştir. Buna göre, Şekil 2’deki son iki grafik, birinci ve ikinci sıra farklarının normal olasılık grafiklerini göstermektedir. Her iki grafikte de görüldüğü gibi, normal dağılım kabul sınırları dışında kalan hiçbir nokta yoktur. Sonuç olarak, her iki sıra için elde edilen farkların yaklaşık normal dağılım gösterdiği söylenebilir.

6. SONUÇLAR

Bu araştırmada astım hastalığına yakalanan çocukların akciğer fonksiyonları üzerinde etkili olan iki farklı denemenin karşılaştırılması amaçlanmıştır. Bunun için, 2x2 çapraz tasarım denemesi yapılmıştır. Denemenin gerçekleştirilmesinden sonra elde edilen verilerin analizi için iki yol seçilmiştir. İlk olarak, aktarılmış etkilerin ve periyot etkisinin olmadığı varsayımı altında verilerin analizi yapılmıştır. Bu analiz sonucunda, iki denemenin hastalar üzerindeki ortalama etkilerinin birbirlerinden önemli düzeyde farklı olduğu ve formoterol ilacının hastalar üzerindeki etkisinin salbutamole göre daha yüksek olduğu görülmüştür. Daha sonra çapraz tasarım analizi gerçekleştirilmiştir. Analiz sonuçları Çizelge 5’te görüldüğü gibidir. Çizelgede F parametresine ait güven aralığı, T değeri ve olasılık düzeyi incelendiğinde, iki denemenin ortalama etkilerinin 5% önem düzeyinde birbirlerinden önemli derecede farklılık gösterdiği anlaşılmıştır. Ayrıca güven aralığının her iki sınır değerinin de negatif olmasından dolayı yine 1 doz formoterolün ortalama etkisinin 1 doz salbutamolden daha fazla olduğu sonucuna ulaşılmıştır. Ayrıca, analiz sonucunda tasarımda aktarılmış etkilerin ve periyot etkisinin önemli olmadığı görülmüştür.

Analiz sonuçları iki denemenin ortalama etkilerinin farklı olduğunu gösterdiğinden biyolojik eşitlik testi yapma zorunluluğu kalmamıştır. Ancak, sonuçlarının tutarlılığının kontrolü amacıyla biyolojik eşitlik testi iki farklı test tipi kullanılarak gerçekleştirilmiş ve beklendiği gibi iki farklı denemenin eşdeğer etkilere sahip olmadığı sonucuna ulaşılmıştır.

Yapılan tüm bu analizler sonucunda, yeni geliştirilmiş formoterol adlı ilacın halen kullanılmakta olan salbutamole göre astım hastaları üzerinde daha olumlu sonuçlar verdiği,

formoterol kullanan astım hastalarının daha kolay solup alıp verebildiği ve birim zamanda soluk verme akışlarının da daha fazla olduğu görülmüştür.

KAYNAKLAR

- Bose M. (2002): "Crossover Designs: Analysis and Optimality Using the Calculus for Factorial Arrangements", Kokata, Design Workshop Lec. Notes ISI, 25-29, s. 183-192.
- Brown B. (1980): "The Crossover Experiment for Clinical Trials", Biometrics, 36, s. 69-79.
- Carriere K.C. (1994): "Crossover designs for clinical trials", Statist. Med., 13, 1063-1069.
- Carriere K.C., Huang R. (2000): "Crossover Designs for Two-Treatment Clinical Trials", Canada, Journal of Statistical Planning and Inference, 87, s. 125-134.
- Carriere K.C., Reinsel G. (1992): "Investigation of Dual-balanced Crossover Designs for Two Treatments", Biometrics, 48, 1157-1164.
- Chinchilli V.M. (2003): "Design and Analysis of Crossover Studies", Hershey, Department of Health Evaluation Sciences, s. 4-20.
- Chow S.C., Liu J.P. (1999): "Design and Analysis of Bioavailability and Bioequivalence Studies", New York, Marcel Dekker Inc.
- Cochran W.G. (1939): "Long-term Agricultural Experiments", J. Royal Statist. Soc.6B, s. 104-148.
- Cochran W.G., Autrey K.M., Canon C.Y. (1941): "A Double Change-over Designs for Dairy Cattle Feeding Experiments", J. Dairy Sci. 24, s. 937-951.
- Donev A.N. (1998): "Crossover Designs with Correlated Observations", Journal of Biopharmaceutical Statistics, 8, 249-262.
- Finney D.J.H. (1965): "Crossover Designs in Bioassay", Proc. Roy. Soc. 145B, s. 42-61.
- Greiner J.M., Johnson W.D. (1993): "Analysis of Crossover Designs with Multi-variate Response", Statistics in Medicine, 12, 69-89.
- Grizzle J.E. (1965): "The Two-period Change-over Design and its Use in Clinical Trials", Biometrics, 21, s. 467-480.
- Hinkelmann K., Kempthorne O. (1994): "Design and Analysis of Experiments Volume I: Introduction to Experimental Design", Canada, John Wiley and Sons Inc., s. 53-59.
- Jones B., Kenward M.J. (1989): "Design and Analysis of Cross-over Trials", New York, Chapman and Hall.
- Keppel G. (1973): "Design and Analysis: A Researcher's Handbook", Prentice-Hall, Englewood Cliffs, NJ.
- Kuehl R.O. (2000): "Design of Experiments: Statistical Principles of Research Design and Analysis", USA, Duxbury Press at Brooks/Cole Publishing Company. Sec. Ed., s. 490-549.
- Mielke P.W.Jr. (1974): "Square Rank Test Appropriate to Weather Modification Crossover Design", Technometrics, 16, s. 13-16.
- Montgomery D.C. (1991): "Design and Analysis of Experiments", Singapore, John Wiley and Sons Inc. Third Ed., s. 1-13.
- Petersen R.G. (1985): "Design and Analysis of Experiments", New York, Marcel Dekker Inc.
- Pocock S.J. (1983): "Clinical Trials: A Practical Approach", East Kilbride, John Wiley and Sons Inc., s. 123-130.
- Senn S. (1993): "Crossover Trials in Clinical Research", England, John Wiley and Sons Inc., s. 6-37.
- Westlake W.J. (1981): "Bioequivalence Testing-a Need to Rethink", Biometrics, Vol. 37, s. 591-593.