

T.C.
MERSİN ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

**KONTROLLÜ RASGELE KLİNİK DENEMELERDE
GRUP ARDIŞIK DENEME DÜZENLERİ TESTLERİNİN
KARŞILAŞTIRILMALI OLARAK İNCELENMESİ**

Mehmet Ali SUNGUR

YÜKSEK LİSANS TEZİ

DANIŞMAN

Doç. Dr. E. Arzu KANIK

MERSİN, 2007

T.C.
MERSİN ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI

**KONTROLLÜ RASGELE KLİNİK DENEMELERDE
GRUP ARDIŞIK DENEME DÜZENLERİ TESTLERİNİN
KARŞILAŞTIRILMALI OLARAK İNCELENMESİ**

Mehmet Ali SUNGUR

YÜKSEK LİSANS TEZİ

DANIŞMAN

Doç. Dr. E. Arzu KANIK

Tez No: 97

MERSİN, 2007

Mersin Üniversitesi Sağlık Bilimleri Enstitüsü

Biyostatistik Anabilim Dalı Yüksek Lisans Programı çerçevesinde yürütülmüş olan “Kontrollü Rasgele Klinik Denemelerde Grup Ardışık Deneme Düzenleri Testlerinin Karşılaştırılmalı Olarak İncelenmesi” başlıklı çalışma, jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 27/06/2007



Doç. Dr. E. Arzu KANIK

Mersin Üniversitesi

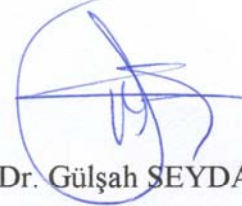
Jüri Başkanı



Yrd. Doç. Dr. Bahar TAŞDELEN

Mersin Üniversitesi

Jüri Üyesi



Doç. Dr. Gülşah SEYDAOĞLU

Çukurova Üniversitesi

Jüri Üyesi

Bu tez, Enstitü Yönetim Kurulunun 11/07/2007 tarih ve 2007/199 sayılı kararı ile kabul edilmiştir.



Doç. Dr. Ülke ÇÖMELEKOĞLU

Enstitü Müdürü



TEŞEKKÜR

Tezimi hazırlarken çalışmalarımın her aşamasında, bilimsel desteğini esirgemeyen, yakın ilgi ve önerileri ile beni sürekli yönlendiren, tüm öğrenim hayatım ve çalışmalarım da tartışılmaz katkıları olan Anabilim Dalı Başkanı ve danışman hocam Doç. Dr. E. Arzu Kanık'a;

Çalışmalarımın en başından sonuna kadar, bilimsel ve manevi anlamda bana sürekli destek olan, karşıma çıkan her konuda yardımlarını esirgemeyen, birlikte çalışmaktan büyük keyif duyduğum değerli hocam Yrd. Doç. Dr. Bahar Taşdelen'e;

Çalışmama gösterdiği hassasiyet, yapıcı ve geliştirici önerileriyle tezime ve akademik hayatıma katkılarından dolayı Çukurova Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı öğretim üyesi Doç. Dr. Gülşah Seydaoğlu'na;

Bütün öğrenim hayatımı borçlu olduğum, beni bugünlere getiren anneme ve babama, tezimi hazırlarken her türlü fedakarlığı göstererek karşılaştığım zorlukları atlatmamda yardımcı olan kardeşlerime

Teşekkür ederim.

Mehmet Ali SUNGUR

Mersin, 2007

Canım Annem'e

İÇİNDEKİLER

KABUL VE ONAY	ii
TEŞEKKÜR	iii
İÇİNDEKİLER	v
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	ix
ÖZET	xi
ABSTRACT	xii
1. GİRİŞ	1
2. GENEL BİLGİLER	5
2.1. KLİNİK DENEMELERE GİRİŞ	5
2.1.1. Klinik Denemelerin Tarihsel Gelişimi	5
2.1.2. Klinik Deneme Nedir?	5
2.1.3. Klinik Denemelerin Aşamaları	6
2.1.3.1.Faz I	6
2.1.3.2.Faz II	7
2.1.3.3.Faz III	7
2.1.3.4.Faz IV	7
2.1.4. Klinik Denemelerin Amaçları	8
2.2. KLİNİK DENEMELERDE DENEKLERİN SEÇİMİ	9
2.2.1. Hedef Populasyon ve Hasta Seçimi	9
2.2.1.1.Uygunluk Kriterleri	9
2.2.1.2.Çalışma Merkezi ve Hasta Seçim İşlemi	10
2.2.1.3.Etik Olarak Dikkate Alınması Gereken Konular	11
2.2.2. Kontrol Grubu ve Seçimi	12
2.3. KLİNİK DENEMELERDE TEMEL İSTATİSTİKSEL KAVRAMLAR	15
2.3.1. Denemelerin Planlanması	15
2.3.2. Etkinlik ve Güvenlik Değerlendirmesi	15
2.3.3. Örneklem Genişliğinin Belirlenmesi	16

2.3.4. Ara (Interim) Analizler ve Veri İzleme (Monitoring) İşlemi	17
2.3.5. İstatistiksel ve Klinik Yorumlama	19
2.3.6. Yanlılık ve Kontrol Altında Tutulması	19
2.3.7. Etki Karışımı	21
2.3.8. Rasgelelik	22
2.3.9. Körleme	24
2.4. KLİNİK DENEMELERDE DENEME DÜZENLERİ	26
2.4.1. Klinik Deneme Düzenleri	26
2.4.2. Klinik Denemelerin Sınıflandırılması	27
2.5. ARDIŞIK DENEME DÜZENLERİ	29
2.5.1. Ardışık Deneme Düzenlerinin Tarihsel Gelişimi	29
2.5.2. Neden Ardışık Deneme Düzenleri	31
2.5.3. Temel Ardışık Deneme Düzenleri	33
2.5.3.1. Tümüyle Ardışık Deneme Düzenleri	33
2.5.3.2. Grup Ardışık Deneme Düzenleri	34
2.5.3.3. Esnek Ardışık Deneme Düzenleri	35
2.6. GRUP ARDIŞIK DENEME DÜZENLERİ	36
2.6.1. İki Muameleyi Karşılaştırmak için Çift Taraflı Hipotez Testleri	36
2.6.2. Sabit Örneklem Genişliği Testleri	37
2.6.3. Grup Ardışık Deneme Düzeni Testleri	38
2.6.3.1. Pocock Testi	39
2.6.3.2. O'Brien & Fleming Testi	42
2.6.3.3. Wang & Tsatis Testi	45
2.6.3.4. Haybittle-Peto Testi	47
3. GEREÇ VE YÖNTEM	50
4. BULGULAR	54
5. TARTIŞMA	77
6. SONUÇLAR VE ÖNERİLER	104
KAYNAKLAR	108
ÖZGEÇMİŞ	111

ŞEKİLLER DİZİNİ

Şekil 2.1	Toplam 5 Analiz İçeren Pocock Testi Sınırları	41
Şekil 2.2	Toplam 5 Analiz İçeren O'Brien & Fleming Testi Sınırları	44
Şekil 5.1	$\alpha = 0,05$ ve $K = 5$ için Kritik Değerler	81
Şekil 5.2	$\alpha = 0,05$ ve $K = 10$ için Kritik Değerler	81
Şekil 5.3	$\alpha = 0,05$ ve $K = 15$ için Kritik Değerler	82
Şekil 5.4	$\alpha = 0,05$ ve $K = 20$ için Kritik Değerler	82
Şekil 5.5	$\alpha = 0,01$ ve $K = 5$ için Kritik Değerler	85
Şekil 5.6	$\alpha = 0,01$ ve $K = 10$ için Kritik Değerler	85
Şekil 5.7	$\alpha = 0,01$ ve $K = 15$ için Kritik Değerler	86
Şekil 5.8	$\alpha = 0,01$ ve $K = 20$ için Kritik Değerler	86
Şekil 5.9	$\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 0,5$ için Örneklem Genişlikleri	90
Şekil 5.10	$\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 1,0$ için Örneklem Genişlikleri	90
Şekil 5.11	$\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 1,5$ için Örneklem Genişlikleri	91
Şekil 5.12	$\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 2,0$ için Örneklem Genişlikleri	91
Şekil 5.13	$\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 2,5$ için Örneklem Genişlikleri	92
Şekil 5.14	$\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 3,0$ için Örneklem Genişlikleri	92
Şekil 5.15	$\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 0,5$ için Örneklem Genişlikleri	93
Şekil 5.16	$\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 1,0$ için Örneklem Genişlikleri	93
Şekil 5.17	$\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 1,5$ için Örneklem Genişlikleri	94
Şekil 5.18	$\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 2,0$ için Örneklem Genişlikleri	94
Şekil 5.19	$\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 2,5$ için Örneklem Genişlikleri	95
Şekil 5.20	$\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 3,0$ için Örneklem Genişlikleri	95
Şekil 5.21	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 0,5$ için Örneklem Genişlikleri	96
Şekil 5.22	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 1,0$ için Örneklem Genişlikleri	97
Şekil 5.23	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 1,5$ için Örneklem Genişlikleri	97

Şekil 5.24	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 2,0$ için Örneklem Genişlikleri	98
Şekil 5.25	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 2,5$ için Örneklem Genişlikleri	98
Şekil 5.26	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 3,0$ için Örneklem Genişlikleri	99
Şekil 5.27	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 0,5$ için Örneklem Genişlikleri	99
Şekil 5.28	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 1,0$ için Örneklem Genişlikleri	100
Şekil 5.29	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 1,5$ için Örneklem Genişlikleri	100
Şekil 5.30	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 2,0$ için Örneklem Genişlikleri	101
Şekil 5.31	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 2,5$ için Örneklem Genişlikleri	101
Şekil 5.32	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 3,0$ için Örneklem Genişlikleri	102

ÇİZELGELER DİZİNİ

Çizelge 2.1 Pocock Testi $C_p(K, \alpha)$ Kritik Değerleri	40
Çizelge 2.2 Pocock Testi $R_p(K, \alpha, \beta)$ Değerleri	41
Çizelge 2.3 O'Brien & Fleming Testi $C_B(K, \alpha)$ Kritik Değerleri	43
Çizelge 2.4 O'Brien & Fleming Testi $R_B(K, \alpha, \beta)$ Değerleri	44
Çizelge 2.5 Wang & Tsatis Testi $C_{WT}(K, \alpha, \Delta)$ Kritik Değerleri	46
Çizelge 2.6 Wang & Tsatis Testi $R_{WT}(K, \alpha, \beta, \Delta)$ Değerleri	47
Çizelge 2.7 Haybittle-Peto Testi $C_{HP}(K, \alpha)$ Kritik Değerleri	49
Çizelge 2.8 Haybittle-Peto Testi $R_{HP}(K, \alpha, \beta)$ Değerleri	49
Çizelge 4.1 $\alpha = 0,05$ ve $K = 5$ için Kritik Değerler	54
Çizelge 4.2 $\alpha = 0,05$ ve $K = 10$ için Kritik Değerler	55
Çizelge 4.3 $\alpha = 0,05$ ve $K = 15$ için Kritik Değerler	55
Çizelge 4.4 $\alpha = 0,05$ ve $K = 20$ için Kritik Değerler	56
Çizelge 4.5 $\alpha = 0,01$ ve $K = 5$ için Kritik Değerler	56
Çizelge 4.6 $\alpha = 0,01$ ve $K = 10$ için Kritik Değerler	57
Çizelge 4.7 $\alpha = 0,01$ ve $K = 15$ için Kritik Değerler	57
Çizelge 4.8 $\alpha = 0,01$ ve $K = 20$ için Kritik Değerler	58
Çizelge 4.9 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 0,5$ için Örneklem Genişlikleri	59
Çizelge 4.10 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 1,0$ için Örneklem Genişlikleri	60
Çizelge 4.11 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 1,5$ için Örneklem Genişlikleri	61
Çizelge 4.12 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 2,0$ için Örneklem Genişlikleri	62
Çizelge 4.13 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 2,5$ için Örneklem Genişlikleri	63
Çizelge 4.14 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 3,0$ için Örneklem Genişlikleri	64
Çizelge 4.15 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 0,5$ için Örneklem Genişlikleri	65
Çizelge 4.16 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 1,0$ için Örneklem Genişlikleri	66
Çizelge 4.17 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 1,5$ için Örneklem Genişlikleri	67

Çizelge 4.18	$\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 2,0$ için Örneklem Genişlikleri	68
Çizelge 4.19	$\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 2,5$ için Örneklem Genişlikleri	69
Çizelge 4.20	$\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 3,0$ için Örneklem Genişlikleri	70
Çizelge 4.21	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 0,5$ için Örneklem Genişlikleri	71
Çizelge 4.22	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 1,0$ için Örneklem Genişlikleri	71
Çizelge 4.23	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 1,5$ için Örneklem Genişlikleri	72
Çizelge 4.24	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 2,0$ için Örneklem Genişlikleri	72
Çizelge 4.25	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 2,5$ için Örneklem Genişlikleri	73
Çizelge 4.26	$\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 3,0$ için Örneklem Genişlikleri	73
Çizelge 4.27	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 0,5$ için Örneklem Genişlikleri	74
Çizelge 4.28	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 1,0$ için Örneklem Genişlikleri	74
Çizelge 4.29	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 1,5$ için Örneklem Genişlikleri	75
Çizelge 4.30	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 2,0$ için Örneklem Genişlikleri	75
Çizelge 4.31	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 2,5$ için Örneklem Genişlikleri	76
Çizelge 4.32	$\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 3,0$ için Örneklem Genişlikleri	76

ÖZET

Kontrollü Rasgele Klinik Denemelerde Grup Ardışık Deneme Düzenleri Testlerinin Karşılaştırılmalı Olarak İncelenmesi

SUNGUR M.A., Klinik denemeler genellikle uzun süreli çalışmalardır ve gerek duyulan deneklerin tamamına aynı anda ulaşmak imkansızdır. Veriler denemenin uygulanması süresince periyodik olarak toplanır. Bu şekilde, denemenin başında belirlenen örneklem genişliğine ulaşmak için yeterli denek bulmak birkaç yıl alabilir. Özellikle ölüm riski veya herhangi bir potansiyel zararı olan denemelerde, bu durum daha da zorlaşabilir ve gerekli örneklem genişliğine ulaşmak deneme süresinin daha fazla uzamasına neden olabilir. Bu nedenle, toplanan verilerin belirli aralıklarla analizlerinin yapılması ve sonuçların değerlendirilmesi araştırmacıların oldukça ilgilendiği önemli bir konudur. Çünkü ara (interim) analizler yapılması ve sonuçları izleme (monitoring) işlemi, yeni muamelenin standart yöntemle göre daha etkin olduğu (superior), etkin olmadığı (inferior) ya da eşdeğer olduğuna (equivalent) ilişkin anlamlı sonuçlara ulaşıldıktan sonra denemenin erken sonlandırılmasını sağlayabilir. Denemenin erken sonlandırılması durumunda, zaman ve kaynaklardan sağlanacak tasarrufun yanında, çalışmaya katılan deneklerin etkin olmayan bir muameleye maruz kalmaları olasılığı daha kısa sürede fark edilerek engellenebilmektedir. Ardışık deneme düzenlerinde, deneme devam ederken toplanan verilere ara analizler yapılır ve bu analizlerin sonuçları sürekli olarak değerlendirilir. Bu tezde ardışık deneme düzenlerine ilişkin bilgiler verildikten sonra grup ardışık deneme düzenleri detaylı olarak tartışılmış ve grup ardışık deneme düzenlerinde uygulanan farklı test tipleri karşılaştırmalı olarak incelenmiştir. Test tiplerinin birbirine göre avantaj ve dezavantajları, gerek duyulan maksimum örneklem genişliği, denemeyi erken sonlandırma potansiyeli ve istenilen test gücü ile daha küçük muamele farklılıklarının tespit edilmesi yönlerinden tartışılmıştır. Sonuç olarak bir grup ardışık deneme planlanırken kullanılacak test tipine karar vermek için çalışılan konunun ne olduğu, kolaylıkla denek bulunup bulunamayacağı ve erken sonlandırma ya da daha küçük muamele farklılıklarını test etme durumlarından hangisiyle ilgilendiği önem kazanmakta ve kullanılacak test tipi bu kriterlere göre farklılık göstermektedir.

Anahtar Kelimeler: Grup Ardışık Deneme Düzenleri, Grup Ardışık Test Tipleri, Ara Analiz, İzleme, VGİK

ABSTRACT

Comparative Analysis of Group Sequential Designs Tests for Controlled Randomized Clinical Trials

SUNGUR M.A., Clinical trials are usually long term studies and it is impossible to reach all required subjects at the same time. Data are accumulated periodically course of the trial. Thus, it may take a few years to enroll enough subjects to meet determined sample size at the beginning of the trial. Particularly, in clinical trials which have death risk or any potential harm this may be more difficult and reaching required sample size may cause more increasing of trial time. Therefore, it is an important interest of investigators substantially to analyze accumulated data in specific intervals and evaluate results. Because, performing interim analyses and monitoring results may provide early termination of trial after obtaining significant results corresponding superiority, inferiority or equivalency of new treatment according to standard method. In condition that early terminating of trial, beside savings provide from time and resources, possibility of subjects participating trial exposure to an inferior treatment may prevent by detecting in a shorter time. In sequential designs, interim analyses are performed to accumulated data and results of these analyses are evaluated consistently as the trial continuing. In this thesis, group sequential designs were discussed in detail after giving some information about sequential designs, and various test types using in group sequential designs were examined comparatively. Advantages and disadvantages of test types according to each other were discussed in points of required maximum sample size, potential to terminating trail early and detecting less treatment differences with desired test power. As a result, for making decision to test type to be used while planning a group sequential trial what will the case be studied, will subjects find easily and which one will the of interest from conditions of early termination or testing less treatment differences are become important, and the test type used varying according to these criteria.

Key Words: Group Sequential Designs, Group Sequential Test Types, Interim Analysis, Monitoring, DSMB

1. GİRİŞ

Klinik denemeler bir ilaç, tedavi veya cerrahi işlem gibi muamelelerin insanların kullanması için güvenli ve etkili olup olmadığını bulmak amacıyla etik kurallar çerçevesinde insanlar üzerinde yürütülen araştırmalardır. Tüm araştırmalarda olduğu gibi, klinik denemelerde de uygun deneme düzenini kullanmak çok önemlidir. Hasta ve kontrol gruplarının belirlenmesi, rasgelelik, körleme gibi konular deneme düzeniyle ilgili olarak dikkate alınması gereken temel noktalarıdır. Çoğunlukla dikkate alınmayan diğer bir önemli konu ise deneme süresince toplanan verilerin izlenmesi (monitoring) işlemidir. İzleme işlemi, öncelikle hastanın yeni muamelenin potansiyel yan etkilerine maruz kalması riskini minimize etmek açısından oldukça önemlidir. İzlemenin diğer bir önemli amacı ise standart yönteme göre yeni muamelenin risk oranı ve yararına ilişkin anlamlı sonuçlara ulaşıldıktan sonra denemenin sonlandırılmasını sağlamasıdır. Yani, yeni bir muamelenin standart yönteme göre daha etkin olduğu (superior) veya etkin olmadığı (inferior) ya da standart yöntemle eşdeğer olduğu (equivalent) anlamlı şekilde saptandıktan sonra, denemenin devam etmesinin gerekliliği yoktur. Denemeleri izleme işlemi karmaşık bir karar işlemidir ve hem istatistiksel metodoloji hem de klinik deneyim gerektirir. İzleme işlemi için genellikle, uygun istatistiksel metodoloji, klinik deneyimleri ve bilimsel kararlarını kullanarak toplanan verileri yeniden incelemeleri için bağımsız bir Veri ve Güvenlik İzleme Kurulu (VGİK) (Data and Safety Monitoring Board (DSMB)) belirlenir (1, 2, 3, 4).

Klinik denemeler, muameleler arasındaki farklılığı tespit etmek için deneme düzeniyle ilgili temel noktalar açısından belirli bir güce (%80, %90 gibi) sahip olabilecek şekilde düzenlenirken, bu farklılığı kabul edilen bazı seviyelerde (%5, %1 gibi) I. tip hata miktarıyla kontrol ederler. Buna uygun olarak iki tip temel deneme düzeni vardır. Bunlar sabit örneklem genişliği deneme düzenleri (fixed sample size designs) ve ardışık deneme düzenleridir (sequential designs) (1). Sabit örneklem genişliği deneme düzenlerinde, tüm veriler toplandıktan sonra sadece bir kere analiz

edilir ve çalışma sonuçları bu analizlere göre yorumlanır. Ardışık deneme düzenlerinde ise deneme süresince toplanan veriler dikkate alınır ve ara (interim) analizler yapılarak toplanan veriler sürekli kontrol edilir. Ardışık deneme düzenleri başlangıçta kalite kontrol alanında ve ekonomik yönden yararları göz önünde bulundurularak geliştirilmiştir. Ürünün başarılı olduğu yönünde bir sonuca ulaşan bir deneme için erken sonlandırma yeni bir ürünün daha çabuk kullanılabilmesi anlamındadır. Ara analizlerde bir pozitif sonuca ulaşılması durumunda, deneme erken sonlandırılarak kaynakların boşa harcanması engellenebileceği gibi ürün denemenin bitmesi planlanan zamandan daha önce piyasaya sürülerek daha çabuk kazanç sağlamaya başlamak da mümkün olabilecektir. Ardışık deneme düzenleri standart sabit örneklem genişliği deneme düzenleri ile karşılaştırıldığında tipik olarak ortaya çıkan tasarruf sağlanabilecek temel noktalar örneklem genişliği, zaman ve maliyet olacaktır. Bunun yanında, denemenin planlandığı şekilde yürütülüp yürütülmediği, deneklerin doğru popülasyondan olup olmadığı, deneklerin uygunluk kriterleriyle uyumlu olup olmadığı ve muamelelerin deneme protokolünde tanımlandığı şekilde uygulanıp uygulanmadığı gibi konuların kontrolü de sağlanabilmektedir. Denemenin başında fark edilmeyen ya da fark edilemeyecek olan bazı problemlerin varlığının tespiti ve çok fazla masraf yapılmadan önce düzeltilmesi de yine ara analizler sayesinde mümkün olacaktır.

Deneklerin insanlar olduğu klinik denemelerde, bireylerin güvensiz, etkin olmayan veya eşdeğer muamelelere maruz kalmalarını engellemek için sonuçların izlenmesi etik bir gerekliliktir. Ayrıca, test edilmek için bekleyen ümit verici başka muameleler de olduğunda bir denemeyi mümkün olduğu kadar çabuk sonlandırmak da etik bir zorunluluktur. Özellikle ölüm riski veya geri dönüşümsüz zararlarla ilgili olan veya herhangi bir potansiyel riski olan klinik denemelerde, sabit örneklem genişliği deneme düzenlerinin kullanılması uygulanabilir olmayabilir ya da kabul edilemeyebilir. Bu gibi klinik denemelerde toplanan verilerin ara analizlerini yapmak etik olarak zorunludur ve izleme işlemi ve VGİK olması da bir zorunluluk olarak dikkate alınmaktadır. Böyle denemelerdeki potansiyel riskleri daha aza indirebilmek amacıyla daha etkin olma (superiority), etkin olmama (inferiority) veya eşdeğerlik (equivalency) sonuçlarından birini elde ederek denemeyi erken sonlandırmaya izin veren ardışık deneme düzenleri geliştirilmiştir. Bu özelliğinin yanında ardışık deneme düzenleri

zaman, maliyet, iş gücü ve uygulama problemleri gibi tüm denemelerde karşılaşılabilecek olan temel sorunların da en aza indirilmesinde etkin bir rol oynamaktadır. Ayrıca deneme süresince yapılan ara analizlerde kurulan deneme düzeninin uygun olmadığı sonucuna ulaşılabilir ve gerekli görülmesi durumunda daha fazla boşa zaman kaybetmeden ilk olarak hedeflenen denek sayısına ulaşılmadan önce deneme düzeni kısmen değiştirilebilir ya da tamamen değiştirilerek yeni bir deneme düzeni kurulabilir. Bir denemeyi sonlandırmaya veya devam etmeye karar vermek için birçok neden olabilmektedir. Tıbbi çalışmalarda, bu nedenler öncelikle yan etkiler, hastaların yaşam kalitesi, denemenin başlangıcında bilinmeyen başka ümit verici muamelelerin de varlığı ve mali giderler olarak sayılabilir. En önemli neden ise, eğer yeni muamelenin standart muameleden daha etkin olduğu veya etkin olmadığı yönünde sonuçlar ortaya çıkar ve bu test istatistikleri kritik değerleri geçerse, VGİK'nun hastalara gereksiz yere etkin olmayan bir muamele uygulanmaması için denemeyi sonlandırmak istemesidir.

Ardışık deneme düzenleri, genel olarak tümüyle ardışık deneme düzenleri (fully sequential designs), grup ardışık deneme düzenleri (group sequential designs) ve esnek ardışık deneme düzenleri (flexible sequential designs) olarak sınıflandırılmaktadır (1). Grup ardışık deneme düzenlerinin en önemli özelliği, ara analizlerin periyodik olarak yapılması ve inceleme için VGİK'na sunulabilmesidir. Grup ardışık deneme düzenleri ortalamaları, oranları veya hayatta kalma eğrilerini karşılaştıran yaygın test istatistikleri için geliştirilmiştir. Grup ardışık deneme düzenleri, ara analizlerin sayısının önceden belirlenip sabit kalmasını ve denemeye giren deneklerin sayısı bakımından eşit aralıklarla yapılmasını gerektirmektedir (1).

Bu tezde, sabit örneklem genişliği deneme düzenlerine bir alternatif olarak ortaya çıkan ardışık deneme düzenlerinden grup ardışık deneme düzenleri ele alınmıştır. Grup ardışık deneme düzenlerinin avantaj ve dezavantajları, sabit örneklem genişliği deneme düzenleri ile kıyaslanarak incelenmiştir. Grup ardışık deneme düzeni uygulanan çalışmalarda kullanılan dört temel test tipi birbirleriyle ve sabit örneklem genişliği deneme düzeni testiyle karşılaştırılmıştır. Karşılaştırmalar, iki muamele ortalamasının karşılaştırıldığı çift taraflı hipotez testleri için yapılmıştır ve aynı I. tip hata (α), test

gücü $(1 - \beta)$ ve analiz grubu sayısı (K) kořulları altında karşılaştırılan iki muamele arasındaki bilinen bir varyans ile normal dağılım gösteren belirli bir farklılığı tespit etme konusunda test tiplerinin birbirlerine göre durumları incelenmiştir. Karşılařtırmalarda testlerin belirli bir I. tip hata (α) düzeyinde H_0 hipotezini reddetmek için kullandığı kritik deęerler, belirli bir test gücünü $(1 - \beta)$ sağlayabilmek için gerek duydukları maksimum örneklem genişlikleri ve denemeyi erken sonlandırarak sabit örneklem genişliği deneme düzeni testlerine göre daha az denek kullanma başarıları incelenmiştir. Test tiplerinin birbirine göre avantaj ve dezavantajları, aynı kořullar altında gerek duyulan maksimum örneklem genişliği, denemeyi erken sonlandırma potansiyeli ve istenilen test gücü ile daha küçük muamele farklılıklarının tespit edilmesi yönlerinden ele alınmıştır.

2. GENEL BİLGİLER

2.1. KLİNİK DENEMELERE GİRİŞ

2.1.1. Klinik Denemelerin Tarihsel Gelişimi

İlk çağlardan bu yana çeşitli muamelelerin bir tedavi yöntemi olup olmadığını anlamak, insanlara acı veren hastalıklara çare bulmak peşinden koşulmuştur. Modern ve gelişmiş bilim ve teknolojinin patlaması, yeni ilaçlar ve yeni tedavi yöntemleri gibi ümit verici başarılı muamele buluşlarını da beraberinde getirmiştir. Bu yeni buluş ve ilaçların klinik araştırmalarına şiddetle ihtiyaç duyulmuştur. Buna paralel olarak, bulunan muamelelerin etkili ve güvenli olduğunu sağlamak için farklı düzenlemeler yapılmış ve uyulması zorunlu olan kanunlar ortaya çıkmıştır. Muamelelerin değerlendirilmesi ve onaylanmasında düzenlemeler koymanın amacı, insan denekler için, özellikle etkinliği ve geçerliliği hala bilinmeyen veya hala araştırma altında olan muamelelerden doğabilecek potansiyel riski minimize etmektir (2).

2.1.2. Klinik Deneme Nedir?

Klinik kelimesi, Yunanca “yatak” anlamına gelen “kline” kelimesinden türetilmiştir (5). Bugünkü terimlerle klinik deneme insanların deneysel birimler olduğu her çalışmadır. Klinik denemeler; test muamelesiyle muamele edilen bir grup hastadaki sonuçları, kontrol muamelesi alan bir grup hastadaki sonuçlarla (her iki gruptaki hastalar da aynı zaman periyodunda kaydedilmiş, muamele edilmiş ve izlenmiş olmalıdır) karşılaştırarak bir veya birden fazla tedavinin etkinliğini değerlendirmek ve insanlara uygulanmak üzere planlanmış deneyler olarak tanımlanabilir (2, 6). Klinik denemenin tanımında ortaya çıkan üç anahtar kelime “deneysel birim”, “muamele” ve

“**muamelenin değerlendirilmesi**”dir. Deneysel birim (denek); çalışmanın sonuçlarının genelleneceği, hedeflenen çalışma popülasyonundan seçilen bir birey olarak ifade edilir. Muamele; genellikle karşılaştırma için referans olarak bir kontrol grubunun bulunduğu, insan denekler üzerinde değerlendirilecek olan bir müdahaledir. Klinik denemelerde muamele; plasebo veya yeni bir farmasötik ilaç veya karışım, tedavi yöntemi, diyet, cerrahi yöntem, tanısal test, tıbbi cihaz ya da sağlık eğitim programının herhangi bir kombinasyonları olabilir.

2.1.3. Klinik Denemelerin Aşamaları

Yeni bir muamele insanlar üzerinde denenmeden önce, güvenli olup olmadığını belirlemek için laboratuvar ve model çalışmalar yoluyla derinlemesine test edilir. Laboratuvar ve hayvan çalışmalarında başarılı bulunduktan sonra kullanıma kabul edilmeden önce insan deneklerde test etmenin 4 aşaması vardır (7, 8, 9, 10, 11). Eğer yeni muamele ilk üç fazı başarılı şekilde geçerse genellikle popülasyonda kullanım için onaylanmaktadır. Bazı muamele tiplerinin popülasyonda kullanım için onaylanmasından önce tüm fazların uygulanması gerekli olmayabilir (10).

2.1.3.1.Faz I

Bir muamelenin insanlarda test edilmesinin ilk aşamasıdır. Yeni muamele, güvenliğini test etmek, muameleyi vermek için en iyi yolun ne olduğunu (tablet, solüsyon, enjeksiyon vb.) bulmak, maksimum tolere edilebilir dozu ve yan etkileri tanımlamak ve farmakokinetik ve farmakodinamiğini değerlendirmek üzere test edilir (2, 3, 11). Amaç, kabul edilemez yan etkilerden sakınacak uygun bir doz düzeyi belirlemektir. Örneklem genellikle 20-80 arasında sağlıklı gönüllüden oluşur. Bu denemeler hemen hemen daima deneğin tam gün tıbbi personel tarafından gözlenebildiği bir klinikte yürütülür.

2.1.3.2.Faz II

Yeni muamelenin başlangıç güvenliği faz I denemelerinde tespit edildikten sonra faz II denemeleri daha büyük gruplarda (100-300 kişi) yapılır ve muamelenin klinik etkinliğini ve güvenliğini değerlendirmek üzere düzenlenir (2, 3, 10, 11). Yani daha büyük gönüllü veya hasta grubunda faz I değerlendirmelerine devam etmek gibidir. Daha fazla hasta muameleyi aldığı anda görülebilecek olan daha az yaygın herhangi bir yan etki olup olmadığı incelenir. Yeni bir muamele için gelişim işlemi yaygın olarak faz II denemeleri süresince zayıf etkinlik veya toksik etkilerin bulunuşu nedeniyle başarısızlığa uğrar (2).

2.1.3.3.Faz III

Büyük hasta gruplarına (1000-3000 veya daha fazla) uygulanan çift-kör kontrollü rasgele klinik denemelerdir. Etkinliğini teyit etmek, yan etkileri izlemek, yaygın olarak kullanılan muamelelerle karşılaştırmak ve muamelenin güvenli şekilde kullanılmasına izin verecek bilgi toplamak için yapılır (2, 3, 9, 11). Bazen 3-5 yıl veya kayıt oranlarına ve gerekli izleme zamanına bağlı olarak daha uzun bir zaman dilimini kapsayabilir. Yeni muamelenin standart muameleden daha iyi, onun kadar iyi veya ondan daha kötü olup olmadığı incelenir. Faz III denemeleri daha karmaşık olabilir ve muamelenin faz I veya faz II denemelerinden daha fazla yönüne bakabilir. Faz III denemeleri, özellikle kronik durumlar için tedavilerde, düzenlemek ve yürütmek için en pahalı, en fazla zaman alan ve en zor denemelerdir.

2.1.3.4.Faz IV

Bir muamelenin başlama sonrası güvenlik gözetimini ve devam eden teknik desteğini kapsar. Bu faz süresince, muamelenin risklerini, yararlarını ve optimal kullanımını kapsayan ek bilgiler araştırılır (2, 3). Bu deneme muamelenin kullanmak için onaylanmasından sonra da olabilir. Denemeler daha iyi doz ayarlama yönergelerini,

yeni formülasyonları, farklı populasyonlar üzerindeki etkiyi veya yeni göstergeleri belirlemek için yürütülebilir (2, 7, 8). Başlama sonrası güvenlik gözetimi, başlangıçtaki klinik denemeler süresince mümkün olandan çok daha büyük hasta populasyonu ve zaman ölçüsü üzerine herhangi bir nadir veya uzun dönem yan etkileri tespit etmek için düzenlenir. Faz IV denemeleri tarafından tespit edilen böyle yan etkiler bir muamelenin kullanımdan kaldırılması veya kısıtlanması ile sonuçlanabilir.

2.1.4. Klinik Denemelerin Amaçları

Klinik denemelerin esas amacı, çalışılan muamele hakkındaki klinik soruları bilimsel olarak ifade etmek için hedef hasta populasyonu hakkında mümkün olan en iyi ve kesinlikle yansız bir yorum elde etmektir. Bir yorumun en iyi kesinliği, bir klinik denemeden elde edilen verilere dayalı olarak tahmin edilen muamele etkisinin değişkenliğinin en küçük olması anlamındadır (2, 6, 11).

Klinik denemelerde ifade edilen önemli klinik/tıbbi soruların uygun ve nihayet bilimsel olarak geçerli bir setinin nasıl geliştirileceği/formülize edileceği, belki de yanlılığı ve değişkenliği minimize edebilmek için yaygın şekilde karşılaşılan zorluklardan biridir. Klinik/tıbbi sorular açık şekilde belirlendikten sonra, bu soruları ifade edecek doğru ve güvenilir bir istatistiksel/klinik yorum yapabilmek için, deneklerin sayısı, çalışma süresi, muameleyi değerlendirmek için çalışma bitim noktaları, olanak/donanım ve klinik personel gibi gerekli kaynakların belirlenmesi gereklidir. Bir klinik denemenin amacı açık, özlu, kesin, bilimsel olarak geçerli ve kolaylıkla hipotezlere çevrilecek durumda kantitatif olan klinik/tıbbi soruların bir seti olarak tanımlanır (2).

2.2. KLİNİK DENEMELERDE DENEKLERİN SEÇİMİ

2.2.1. Hedef Populasyon ve Hasta Seçimi

İstatistiksel/klinik yorum genellikle hedef hasta populasyonundan seçilen temsil edici bir örnekleme dayalı olarak elde edilir. Bu nedenle, klinik bir denemede hastaların seçimi, çalışılan muamele hakkındaki bilimsel ve/veya tıbbi sorulara en iyi yanıt vermede önemli bir noktadır. Temel olarak bir klinik deneme için hastaların seçimi iki aşamalıdır (2). İlk olarak hedef hasta populasyonu tanımlanmalıdır. Daha sonra hastalar bu hedef hasta populasyonundan seçilir. Herhangi bir hastalık için hedef hasta populasyonu, genellikle hasta özellikleri ve hastalığın şiddeti bakımından oldukça heterojendir. Kuşkusuz hedef hasta populasyonunun heterojenliği, çalışmanın doğruluğuna, güvenilirliğine ve çalışma bulgularının genellenmesine olumsuz yönde etki edecektir. Hedef hasta populasyonu genellikle gözlenen klinik sonuçların doğruluğunu ve güvenilirliğini azaltacak çeşitli beklenen ve beklenmeyen yanlılık ve değişkenlik kaynakları içerir (2, 11, 12, 13). Örneğin; yaş, cinsiyet, boy, ağırlık ve fonksiyonel statü gibi demografik özelliklerdeki farklılıklar nedeniyle olan yanlılık ve değişkenlikler beklenenlerdir. Hastalık statüsündeki ve aynı zamanda yapılan tedavilerdeki değişikliklerden kaynaklanan yanlılık ve değişkenlikler ise beklenmeyenlerdir. İyi bir klinik uygulama için, yanlılığı azaltmak ve değişkenliği minimize etmek, dolayısıyla hedef hasta populasyonunu bu özellikler bakımından mümkün olduğunca homojen olarak belirlemek gerekmektedir.

2.2.1.1.Uygunluk Kriterleri

Klinik denemelerde, denemeye girmek üzere uygun hastaların seçilebildiği hedef hasta populasyonunu tanımlamak için genellikle bir **uygunluk kriterleri** seti geliştirilir. Tipik olarak bir uygunluk kriterleri seti, bir **dahil etme kriterleri** seti ve bir **hariç tutma kriterleri** setinden oluşur (2, 11). Dahil etme kriterleri seti hedef hasta populasyonunun özelliklerini kabaca belirlerken hariç tutma kriterleri ise beklenen

değişkenlik kaynaklarını kaldırarak hedef hasta popülasyonunun daha detaylı bir profilini çıkarır. Hastalar, denemeye girebilmek için tüm dahil etme kriterlerine uygun olmak zorundadır. Hariç tutma kriterlerinden herhangi birine sahip olan hastalar ise çalışmadan çıkarılır. Uygunluk kriterleri hasta özellikleri, tanıya yönelik kriterler, muamele süresi ve hastalığın şiddetine dayalı olarak geliştirilmelidir (2). Hasta seçimi için iyi geliştirilmiş bir uygunluk kriterleri seti, tek başına hedef hasta popülasyonunu en iyi şekilde tanımlayamaz, fakat homojen bir örneklem elde edilmesini sağlar. Ayrıca uygunluk kriterleri, yanlılık ve değişkenliğin azalmasına yardım ederek çalışmanın istatistiksel gücünün artmasına da katkıda bulunur.

2.2.1.2.Çalışma Merkezi ve Hasta Seçim İşlemi

Klinik denemelerin, istenilen kesinlik ve doğrulukla çalışılan muamelenin geçerli bir değerlendirmesini sağlaması için, hastaların sayısı genellikle çalışma protokolünde belirtilir. Pratikte, tek bir çalışma merkezi sınırlı kapasitesi ve kaynakları nedeniyle amaçlanan klinik deneme için uygun olmayabilir. Ayrıca o merkezde, çalışmanın planlanan zaman periyodu içinde, mevcut hastalığa sahip yeterli hasta olmayabilir. Yeterli sayıda hasta kaydedebilmek ve çalışmayı belirlenen zaman süreci içerisinde tamamlayabilmek için bir alternatif olarak çok merkezli denemeler planlanır (14). Çok merkezli bir deneme kurulduğunda, kaç merkez kullanılacağı ve bu merkezlerin nasıl seçileceği önemlidir. Bir kural olarak, merkezlerin sayısı seçilen her bir merkezdeki hasta sayısından büyük olmamalıdır (2, 14). Çünkü muameleler arasındaki istatistiksel karşılaştırma genellikle çalışma merkezlerindeki hastalara dayalı olarak yapılır. Çalışma merkezlerinin seçimi aşağıdaki kriterlere dayalıdır (2, 14):

- Araştırmacının çalışılan hastalıkla ilgili bireysel kalifiyesi ve deneyimi
- Amaçlanan denemeyi kurmak için araştırma merkezinin uygunluğu
- Araştırma merkezindeki personelin eğitimi, öğretimi ve deneyimi
- Gerekli araçların (görüntüleme cihazları, laboratuvar analiz cihazları vb.) varlığı
- Coğrafik konum

Bu kriterlere dayalı olarak çalışma merkezleri belirlendikten sonra, seçilen çalışma merkezlerinde hastaların seçim işlemleri şu kriterlere göre yapılır (2):

- Kaç hastanın uygunluk kriterleriyle uyumlu olduğunun baştaki tahmini
- Tanısal kriterlere dayalı görüşme
- Hastaların hastalık değişimleri
- Eş zamanlı hastalık/tıbbi tedavi
- Psikolojik faktörler
- Aydınlatılmış onam

2.2.1.3.Etik Olarak Dikkate Alınması Gereken Konular

Şiddetli, tahrip edici veya ölümcül çoğu hastalıklar (AIDS, Alzheimer, kanser vb.) için etkinliği bilinen bir alternatif muamele mevcutken bir klinik denemede eşzamanlı plasebo kontrolünün kullanılması etik değildir. Bununla birlikte, bir muamelenin etkinlik ve güvenliğini test etmek için yapılan bir deneme, sadece eşzamanlı plasebo kontrolü kullanarak kurulabilmektedir. Etik dikkate almalar kesinlikle hasta seçim işlemini etkilemektedir. Böyle durumlarda, muamele kolu ve plasebo koluna, plasebo koluna atanan hastaların oranı daha az olacak şekilde farklı hasta sayılarının atanması önerilmektedir (2).

Çocuklar, yaşlılar ve hamile veya hamilelik potansiyeli olan bayanların klinik denemelere dahil edilmesi önemli bir konudur. Yeni doğanlar, bebekler ve çocuklar belirli tedavilere yetişkinlerden farklı şekilde yanıt verdiği için, bu yaş gruplarında kullanılacak tedavilerin klinik denemeleri daima gereklidir. Bu nedenle, çocuklardaki denemelerin yapısına özel bir önem verilmelidir. Benzer şekilde, yaşlılarda yan etkilerin insidansının artması ve bazı tıbbi tedavilerin farklı farmakokinetik mekanizmaları ve bozulmuş homeostaz gibi nedenlerden dolayı bu yaş grubunda kullanılma olasılığı olan herhangi bir tıbbi tedavinin değerlendirilmesi de oldukça önemlidir (2).

2.2.2. Kontrol Grubu ve Seçimi

Çalışma sonunda yararlı veriler elde edilemeyeceğinden ve sonuçları yorumlamak zor olacağından kontrol grubu olmayan (karşılaştırmaz) çalışmaların kullanılması önerilmemektedir. Etkinliğin yansız ve geçerli bir değerlendirmesini ve çalışmanın güvenliğini sağlamak için, uygun kontrollü klinik denemeler planlanmalıdır. Kontrollü bir çalışmanın amacı sadece yanlılığı elemek değil, ayrıca çalışmanın değişkenliğini minimize etmek ve buna bağlı olarak istatistiksel ve klinik yorumunun doğruluğunu ve güvenilirliğini artırmaktır. Başka bir deyişle, kontrol grubu olmayan çalışmalar, test edilen muameleye tabi tutulmamış olsalardı hastalara ne olacağı hakkında yansız ve güvenilir bir yoruma sahip olamazlar. Kontrollü bir çalışma, gruplarda yer alan hastaların çalışmada beklenen sonuçları veya klinik bitim noktalarını etkileyebilecek olan tüm demografik ve temel özellikleri ile deneme sırasında muamele aşamasındaki tüm değerlendirmeler ve değişkenler bakımından muamele grupları arasında homojenlik olacak şekilde planlanmış bir deneme olarak tanımlanabilir (2, 15). Kontrol grupları iki kritik özelliğe göre sınıflanabilir: Kimin kontrol grubunda olacağını belirleme metodu ve hastaların alacağı muamele tipi. Hastaların kontrol grubuna atanması, ya aynı hedef hasta popülasyondan rasgele şekilde hastaların seçilmesi ya da denemede muamele edilecek olan hedef hasta popülasyonundan ayrı bir kontrol hasta popülasyonu belirlenmesiyle olabilir (2).

Genellikle kontrol grubundaki bir hastanın alabileceği dört muamele tipi vardır: plasebo, muamele yok, test muamelesinin farklı uygulaması (farklı ilaç dozu) veya farklı bir aktif muamele. Buna göre bir klinik denemede kullanılacak kontrol grupları şu şekilde sınıflandırılabilir (2, 15):

1. Eşzamanlı Plasebo Kontrol (Placebo Concurrent Control)
2. Eşzamanlı Doz-Yanıt Kontrol (Dose-Response Concurrent Control)
3. Eşzamanlı Aktif (Pozitif) Kontrol (Active (Positive) Concurrent Control)
4. Eşzamanlı Muamele Olmayan Kontrol (No Treatment Concurrent Control)
5. Harici (Tarihsel) Kontrol (External (Historical) Control)

Plasebo, tıbbi tedavinin bir şekli veya tıbbi tedaviyi taklit etmek için uygulanan bir müdahaledir (2). Genellikle muamelenin sembolik etkisini görmek veya kontrollü bir klinik denemede bilgi toplama yanlılığını (information/observation bias) elemek için kullanılır. Klinik denemelerde plasebo etkisini gözlemek oldukça yaygındır. Bazı durumlarda plasebo etkisini gözlemek çoğu muamelenin etkisini gözlemek kadar önemli olabilmektedir. Çalışılan tedavisel muamelenin etkinlik ve güvenliğinin net bir şekilde ve tamamen yansız bir değerlendirmesinin sağlanması için, klinik denemelerde eşzamanlı plasebo kontrolünün kapsanması gereklidir. Bununla birlikte, şiddetli veya ölümcül hastalıklarla ilgili olan klinik denemelerde, özellikle de etkinliği kanıtlanmış olan alternatif bir muamele mevcut olduğunda eşzamanlı plasebo kontrolünün kullanılması etik değildir. Bunun dışında kalan ve uygulanması mümkün olan tüm durumlarda ise, yeni bir muamelenin etkinlik ve güvenliğinin değerlendirilmesinde standart eşzamanlı kontrol grubu olarak plasebo eşzamanlı kontrol kullanılmalıdır (2).

Eşzamanlı doz-yanıt kontrollü bir klinik deneme, genellikle muamelenin minimum etkin dozunun ve maksimum tolere edilebilir dozunun belirlendiği faz II denemelerde kullanılır ve bu denemelerde aynı muamelenin en az iki farklı dozu kapsar. Faz II denemelerinde henüz muamelenin etkinliği kanıtlanmamış olduğundan doz-yanıt ilişkisi bakımından her bir doz için mutlak etkinliğin saptanabilmesi için eşzamanlı plasebo kontrolün de kapsanması zorunludur (2).

Klinik denemelerde ilgilenilen konu, yeni muamelenin standart muameleye göre daha etkin olduğunu veya daha iyi bir güvenlik profili göstererek eşdeğer bir etkiye sahip olduğunu kanıtlamak olabilir. Bu amaca sahip klinik denemeler genellikle bir eşzamanlı aktif muamele kontrol grubu içererek kurulur (2). Yukarıda ifade edildiği gibi, bazı durumlarda ise eşzamanlı aktif muamele grubu içermek için etik zorunluluklar olabilir. Eğer deneme yeni muamelenin etkinliği ve güvenliği hakkında önemli bir kanıt sağlamak amacıyla kuruluyorsa, eşzamanlı aktif muamelenin daha etkin olduğu, eşzamanlı plasebo kontrolüyle yapılan denemelerde gösterilmiş olmak zorundadır. Aksi takdirde denemede eşzamanlı aktif muamele kontrol grubuna ek olarak eşzamanlı plasebo kontrolü grubunun da bulunması bir zorunluluktur (2).

Tarafsız etkinlik ölçümlerinin mevcut ve kısa bir zaman diliminde elde edilebilir olması ve plasebo etkisinin ihmal edilebilir olması ön şartlarına sahip belirli hastalıklar için, test muamelesi eşzamanlı olarak muamele olmayan bir kontrol grubuyla karşılaştırılabilir (2). Bununla birlikte, pratikte, mümkün olduğu sürece eşzamanlı muamele olmayan kontrol grubuyla çalışmaktan kaçınılması önerilir. Çünkü bu durum, yeni muamelenin faz I, faz II ve faz III denemeleri için iyi bir klinik uygulama değildir ve plasebonun etkinlik üzerine psikolojik etkisini taklit etmekte başarısız olabilir (2).

Klinik denemelerde bazen test muamelesinin sonuçlarını, diğer aktif muamelelerin sonuçlarıyla veya bir hastalığın uygun şekilde belgelenmiş tarihsel deneyimiyle karşılaştırmakla ilgilenilebilir. Tarihi veriler iki şekilde elde edilebilir (2). Farklı zamanlarda hiçbir muamele almamış, aynı muameleyi almış veya farklı muameleler almış aynı hasta grubundan elde edilebilir. Farklı zamanlarda hiçbir muamele almamış, aynı muameleyi almış veya farklı muameleler almış farklı bir hasta grubundan elde edilebilir. Her iki durumda da tarihsel kontrol grubuna ait veriler eşzamanlı değildir ve kontrol grubu ve çalışma grubuna ilişkin veriler eşzamanlı olarak gözlenemez (2).

2.3. KLİNİK DENEMELERDE TEMEL İSTATİSTİKSEL KAVRAMLAR

2.3.1. Denemelerin Planlanması

Klinik denemelerin planlama aşamasında, “sorunun ne olduğu”nun tanımlanması önemlidir, sorunun ne olduğunun tanımlanması çalışmanın amacını belirlemede önemli bir etkidir. Sorunun ne olduğunun tanımlanmasından sonraki önemli nokta ise “sorunun nasıl cevaplanacağı”dır (2). Esas olan amaçlanan klinik denemenin, sorunun yansız ve geçerli bir bilimsel değerlendirmesini sağlamasıdır.

Dikkatli ve iyi organize edilmiş bir protokol çalışmanın amacı, deneme düzeni, hasta seçim kriterleri, doz alım çizelgesi ve istatistiksel metotlar gibi konuları içerir (2, 11). Sonuç olarak, “uygun bir çalışma düzeninin seçilmesi” ve “toplanan verilerin geçerli istatistiksel metotlar kullanılarak analiz edilmesi” bir klinik deneme planının iki önemli noktasıdır. Kullanılacak istatistiksel metotlar, kullanılan deneme düzenine dayalı olduğundan bu iki konu birbiriyle yakından ilgilidir. Bir deneme düzeni seçilmeden önce, istatistiksel analizlerin kullanımını desteklemek amacıyla klinik denemelerin amaçları, hasta seçimi, rasgelelik, körleme, kontrollerin seçimi ve diğer bazı istatistiksel sorunlar gibi deneme düzeniyle ilgili temel konular dikkate alınmalıdır.

2.3.2. Etkinlik ve Güvenlik Değerlendirmesi

Bir klinik denemeyle, konuyla ilgili tüm istatistiksel/klinik soruları ifade etmek mümkün değildir. Bu nedenle, ilgilenilen bilimsel ve tıbbi soruları ifade etmede kullanılacak birincil ve ikincil yanıt değişkenlerini tanımlamak önemlidir. Çalışma amaçlarına ulaşmakta ihtiyaç duyuldukları için yanıt değişkenleri (veya klinik bitim noktaları) genellikle başlangıçta seçilmektedir. Pratikte, seçilen klinik bitim noktalarının geçerliliği kanıtlanmış (güvenilir ve tekrarlanabilir), yaygın olarak mevcut, anlaşılabilir ve kabul edilebilir olması önerilmektedir (2).

Etkinlik deęerlendirmesi iin, birincil etkinlik deęiřkeni tanımlandıktan sonra, hastaların deęerlendirilebilirlięi iin kriterler kesin olarak tanımlanmalıdır. rneęin; herhangi bir etkinlik gzlemi ile veya belirli bir minimum gzlem sayısı ile tm hastalara dayalı olarak bir analiz kurulabilir. Etkinlięin adil bir deęerlendirmesini saęlamak iin, bazen denemenin seyri boyunca sadece belirli bir uyuma kriterine sahip olan hastalar (dozun %80 ve %120'sini alanlar gibi) analize alınabilir. Gvenlik deęerlendirmesi iin muameleye giren ve en azından muamelenin bir dozunu almıř olan tm hastaların kapsanması gerekir. Gvenlik deęerlendirmesi genellikle klinik ve laboratuvar testlerine dayalı olarak yapılır (2).

2.3.3. rneklem Geniřlięinin Belirlenmesi

Muameleler arasında gerekte var olan anlamlı bir farklılıęı doęru řekilde tespit etme olasılıęı istatistiksel olarak **testin gc** řeklinde ifade edilir. Gerekte farklılık yokken muameleler arasında farklılık olduęu sonucuna ulařılması ise **I. tip hata** olarak tanımlanır ve bu hatanın miktarı genel olarak alıřmanın bařında belirlenir. Belirlenen I. tip hata dzeyinde, rneklem geniřlięi artırılarak testin gc de artırılabilir. Bir rneklem geniřlięi belirlemesinin amacı klinik bilim adamları tarafından saęlanan bilgiye (istenilen g, deęiřkenlik ve klinik olarak anlamlı farklılıklar vs.) dayalı olarak uygun bir rneklem geniřlięi hesaplamaktır. rneklem geniřlięinin belirlenmesi, seilen deneme dzeni altında uygun istatistiklere dayalı olarak yapılmalıdır. Farklı deneme dzenleri ve farklı hipotez testleri yaparken, istenilen gce ulařmak iin farklı rneklem geniřlięi gereksinimleri ortaya ıkabilir. Bu nedenle rneklem geniřlięinin belirlenmesi iin g analizi yaparken ařaęıdakilerin dikkate alınması tavsiye edilmektedir (2, 16, 17, 18):

- Hangi deneme dzeninin kullanılacaęı
- Hangi hipotezlerin test edileceęi
- Hangi istatistiklerin hesaplanacaęı

2.3.4. Ara (Interim) Analizler ve Veri İzleme (Monitoring) İşlemi

Ara analizler ve veri izleme işlemi, bir klinik denemede denemenin başlangıcından itibaren toplanan verilerin analiz edilmesi, incelenmesi ve sonuçların izlenmesi işlemleridir. Ara analizlerin varlığı, deneme düzeninin yapısı ve seçilen deneme düzeniyle ilgili olmakla birlikte, bazı denemeler için ara analizlerin yapılması ve sonuçlarının incelenerek muamelenin etkinlik ve güvenliğinin değerlendirilmesi etik bir zorunluluktur. Bir denemede ara analizlerin yapılması planlanırken dikkat edilmesi gereken noktalar şunlardır (2, 4, 19, 20):

- Belirlenen toplam I. tip hata miktarının korunması
- Körlemenin korunması
- Ara analizlerin, gelecek çalışmalara hipotezler üretmek veya güvenilirliği değerlendirmek gibi yönetsel nedenler için veya planlama amaçları için yapılması

Toplam I. tip hata miktarının korunması genellikle dikkatli şekilde planlanmış bir çalışma protokolü ile başarılabilir. Körlemenin korunması ise yapılan çalışmanın güvenilir olması konusunda büyük bir etkiye sahip olan önemli bir konudur. Üçüncü madde ise ara analizleri, denemenin uygulanması süresince karşılaşılabilecek olan bazı önemli soru ya da sorunları belirlemek için yönetsel bir araç olarak dikkate almaktadır.

Ara analizler ve veri izleme işleminin en önemli özelliği, hastalara uygulanan muamelelerde daha etkin olma durumu ya da aşırı bir güvenlik riski gözleendiğinde denemeyi sonlandırmak için bir karar sağlamasıdır. Eğer ara analizlerin yapılması planlanmamışsa, bunun nedenleri çalışma protokolünde belirtilmelidir. Ara analizlerin sonuçlarına göre deneme düzeninde herhangi bir değişikliğe gidilebileceği için ve potansiyel bir yanlılık riski ortaya çıkabileceği için, deneklerin ara analizlerin sonuçları hakkında herhangi bir bilgisi olmamalıdır. Ara analizler yapıldığında bu körlüğü sağlamak için aşağıdaki noktalara dikkat edilmesi oldukça önemlidir (2, 4, 9, 21):

- Ara analizlerin kullanımı için rasgelelik, hastaların denemedeki tanımlama numaralarıyla herhangi bir ilgisi olmayan bir istatistikçi tarafından yapılmalıdır
- Bireylere uygulanan muamele tipi bilinmelidir
- Ara analizler, muamelenin etkisi üzerine karar vermek ve ara analizlerin amaçlarına uyum sağlamak için gerek duyulan sayıda bireyle yapılmalıdır
- Körlük, karar verme ve ara analizlerinin sıklığı gibi detaylı işlemler, harici yeniden incelemeler için belgelenmeli ve bu belge incelenmeye hazır bulundurulmalıdır

Ara analizler, klinik denemeler süresince rasgele dağıtım kodlarının körlenmemiş olması yanında, sonuçların hem harici hem de dahili Veri ve Güvenlik İzleme Kuruluna (VGİK) sunulmasını da gerektirdiğinden, denemenin körlüğü ciddi şekilde tehlikeye atılmaktadır ve potansiyel bir yanlılık daima ortaya çıkacaktır. Bu nedenle, ara analizlerin yapılması ve verilerin izlenmesi işlemleri konusunda, yukarıdaki maddelere ek olarak aşağıdaki noktalara da dikkat edilmesi tavsiye edilmektedir (2):

- Harici VGİK'nda, denemenin yürütücüleri arasında olmayan ve çalışılan hastalık alanından olan akademik enstitülerden klinik uzmanlar ve istatistiksel uzmanlar yer almalıdır
- Ara analizler, VGİK'nun istatistiksel üyeleri arasında olmayan bir istatistik uzmanı tarafından, protokolda ara analiz için önceden belirlenmiş etkinlik ve güvenlik bitim noktalarına göre rasgele dağıtım kodları ve hasta tanımlama numaralarını birleştirilen veritabanını kullanarak yapılmalıdır
- Harici VGİK, ara analizlerin sonuçlarının ne zaman, nasıl ve hangi üst birim ya da kurumlara sunulmak üzere mevcut olması gerektiği hakkında karar verecek yetkiye sahip olmalıdır

2.3.5. İstatistiksel ve Klinik Yorumlama

İstatistiksel ve klinik yorumlamalar genellikle kontrollü rasgele klinik denemelerden toplanan verilere dayalı olarak yapılır. İstatistiksel ve klinik yorumlamalar seçilen örneklemin hedef hasta populasyonunu iyi temsil eden bir örneklem olduğu varsayımı altında istatistiksel testlerden elde edilir. Genellikle rasgele seçilmiş bir örneklem, temsil edici bir örneklem olarak ifade edilir.

İstatistiksel yorumun klinik denemelerden elde edilen verilerle tümevarım işleminin bir parçası olduğu ve anlamlı bir klinik yorumun olasılığını engellememesi gerektiği unutulmamalıdır. Deneme kurulmadan önce, dahil etme ve hariç tutma kriterleri çalışma protokolünde kesin olarak belirlenmişse, örneklemdaki hastaların temel demografik özellikleri, bu örnekleminin çekildiği hasta populasyonunu tanımlamak için kullanılabilir.

Nadiren tek bir denemenin sonuçlarına dayalı olarak klinik yorum yapılır. Bir muamelenin etkinlik ve güvenliğini değerlendirmek için, genel olarak benzer ama farklı özellikleri olan hastalarla, benzer deneme düzeni ve eş zamanlı kontrol kullanan bir kaç klinik deneme kurulur. Bu klinik denemeler farklı ülkelerde, farklı çalışma merkezlerinde ve farklı araştırmacılar tarafından da kurulabilir (2).

2.3.6. Yanlılık ve Kontrol Altında Tutulması

Yanlılık, bir denemenin herhangi bir aşamasında ortaya çıkarak verileri gerçekten saptıran ve doğrudan çalışma sonucunun yanlış bulunmasına neden olan sistematik hata veya hatalar bütünüdür (22). Yanlılık var olan bir ilişkiyi maskeleyebilir veya sahte bir ilişkiye neden olabilir ve etki büyüklüğünün fazla veya az tahminine neden olabilir. Bir denemenin planlanmasındaki veya bir değişkenin ölçümündeki hatayı içeren yanlılık, bir klinik denemenin herhangi bir aşamasında meydana gelebilir ve esas olarak dört kaynaktan ileri gelir:

- Seçilen deneme düzeni
- Deneme süreci
- İstatistik analiz
- Sonuçların yorumu

Çok çeşitli yanlılık tipleri vardır ve genel olarak **Seçim Yanlılığı (Selection Bias)** ve **Bilgi Toplama Yanlılığı (Information/Observation Bias)** şeklinde 2 grupta sınıflandırılmaktadır (2, 12, 13, 23, 24):

Seçim yanlılığı, çalışmaların planlama aşamasında meydana gelir. Genellikle vaka-kontrol çalışmalarında vakaların ve kontrollerin seçimi muameleyle ilgili farklı kriterlere dayandırıldığında ve çalışılan örneklem, hedef popülasyonu temsil etmediğinde ortaya çıkan hatadır (23, 24). Örneklem hedef hasta popülasyonunu iyi temsil etmediğinde etki tahmininde bir hata olur. Seçim yanlılığının klinik denemelerde meydana gelebilen en yaygın yanlılık kaynağı olduğu söylenebilir. Bazen deneklerin bazıları çalışmaya alınıp bazıları alınmadığında, bir denemenin uygulanması aşamasında da meydana gelebilir. Seçim yanlılığına neden olan hata kaynakları genel olarak şunlardır (25):

- Uygun olmayan popülasyonun çalışılması
- Yetersiz katılım
- Ayırıcı sınıflamanın değişimi
- En kolay bulunanlar veya gönüllülerin seçilmesi

Bilgi toplama yanlılığı, çalışmaların veri toplama aşamasında meydana gelir. Ölçümdeki bir hata ya da denekleri muamele ve/veya sonuç değişkenleri için yanlış sınıflandırma nedeniyle tahmin edilen etki çarpıtıldığında ortaya çıkar. Bilgi toplama yanlılığı kategorilerinin her biri içinde belirlenen yanlılıklar, farklılığa ait olan hatalı sınıflandırma belirgin şekilde meydana gelmesine rağmen hem farklılığa ait olan hem de olmayan hatalı sınıflandırma özelliklerine sahip olabilirler (23, 24). Genellikle bilgi toplama yanlılığına neden olan hata kaynakları şunlardır (25):

- Denek bilgisi
- Gözlemci bilgisi
- Araçların kusuru
- Ölçümdeki teknik hatalar

Bazı yanlılıklar belirli bir analitik çalışma tipine özelken bazı yanlılık tipleri ise epidemiyolojik araştırmalarda kullanılan tüm temel deneme düzenlerinde ortaya çıkabilirler. Bir klinik denemenin herhangi bir yanlılıktan bağımsız olması hemen hemen imkansızdır. Bu nedenle, bir deneme planlanırken potansiyel yanlılık kaynaklarını tespit etmek gerekmektedir. Denemenin herhangi bir aşamasında meydana gelebilecek olan herhangi bir potansiyel yanlılığı tanımlamak ve bu yanlılık kaynağından etkilenmeden denemenin nasıl kurulacağına karar vermek çok önemlidir. Potansiyel yanlılık tanımlandıktan sonra bu yanlılığı minimize etmek ya da çalışmadan tamamen elemek için, deneme düzeni seçiminde dikkat edilecek noktaların dışında körleme veya rasgelelik gibi işlemlerin uygulanması da yanlılığın minimize edilmesi ya da tamamen ortadan kaldırılmasında önemli katkılar sağlayacaktır. Buna ek olarak, ideal bir klinik denemedeki hastalar, denemenin sonuçlarından genelleme yapılacak olan hedef hasta popülasyonundan tamamen rasgele seçilmiş ve bu popülasyonu en iyi temsil eden bir örneklem olmalıdır (12, 13).

2.3.7. Etki Karışımı

Tanım olarak etki karışımı; çalışılan muamelenin sonuç üzerine etkisinin, sonuç için bağımsız bir risk faktörü ve muameleyle ilişkili üçüncü bir faktörle karışımıdır (13, 23, 26). Genellikle veri analizi aşamasında meydana gelir ve tahmin edilen ilişkinin doğru etki ile aynı olmamasına neden olur. Doğru ilişkinin fazla ya da düşük tahminine neden olabilir, belki de gözlenen etkinin yönünü tamamen değiştirebilir. Etki karıştırıcılar pozitif veya negatif olabilir. Pozitif etki karıştırıcılar bir ilişkinin fazla tahminine neden olurken, negatif etki karıştırıcılar ise ilişkinin düşük tahminine neden olurlar (23, 24, 26). Etki karışımının klasik örneği, sigara içme, alkol tüketimi ve akciğer kanseri arasındaki ilişkidir. Buradaki etki karıştıran, alkol kullanımıyla ilişkili

ve akciğer kanseri için bir bağımsız risk faktörü olan sigara içmedir. Sigara içenlerde akciğer kanseri riski arttığından, böyle bir denemede sigaradan kaynaklı akciğer kanseri vakaları da alkol kullanımıyla ilişkilendirilebilir ve alkol kullanımının akciğer kanseri üzerindeki etkisi gerçekte olduğundan daha büyük bulunabilir.

Etki karışımı yanlılığa benzerdir ve genellikle de karıştırılır. Bazen yanlılık tipleri arasında bile sayılmaktadır. Ancak yanlılık bir denemenin planlanmasındaki veya bir değişkenin ölçümündeki hatayı içerirken etki karışımı doğru planlanmış bir denemenin veya doğru olabilecek bir ölçümün istatistik modellemede yanlış kullanılmasından dolayı sonuçların hatalı olmasına neden olur ve doğru model geliştirildiğinde hata değil bir bilgi kaynağıdır (22).

2.3.8. Rasgelelik

Uygun ve doğru rasgelelik, seçim yanlılığını eler ve yüksek kalite rasgele klinik denemelerde çok önemli bir unsurdur. Rasgelelik, tüm katılımcılara eşit çalışma veya kontrol muamelesi alma şansı verir. Deneklerin muamele gruplarına atanması birbirinden bağımsız ve yansızdır, deneklerin veya araştırmacının tercihlerine veya önyargılarına yer vermez.

Rasgelelik için temel metotlar şunlardır (7, 8, 9, 27, 28):

1. Basit Rasgelelik (Simple Randomisation)
2. Minimizasyon (Minimisation)
3. Bloke Edilmiş (Sınırlanmış) Rasgelelik (Blocked (Restricted) Randomisation)
4. Tabakalı Rasgelelik (Stratified Randomisation)
5. Küme Rasgelelik (Cluster Randomisation)

Basit Rasgelelik, rasgele sayı tablosu kullanarak, yazı tura atarak veya bilgisayar programları yardımıyla yapılan muamele atamasıdır (27).

Minimizasyon, özellikle küçük örneklerde, önemli özellikler bakımından gruplar arasındaki dengeyi düzenler. Denemeye girecek olan sıradaki katılımcının o aşamada gruplar arasındaki tüm dengesizliği minimize edecek olan muameleye verilmesi düşüncesine dayanır (27, 28).

Bloke edilmiş rasgelelik, blokların içinde denge sağlamak için blokların içinde muamelelerin rasgele atanmasıdır. Bloklar herhangi bir büyüklükte olabilir fakat deneme gruplarının sayısının katı veya katları mantıklıdır (9, 27, 28).

Tabakalı rasgelelik, önemli katılımcı özellikleriyle tanımlanan alt grupların içinde bir denge verir. Tabakalama çok değişkenli küçük çalışmalar için uygun değildir, büyük denemeler için de gerekli değildir (7, 8, 27).

Küme rasgelelikte, rasgelelik birimi çalışılan katılımcı birey değil katılımcıların veya kümelerin gruplarıdır. Bu plan özellikle muamele bir grup düzeyinde olduğunda veya muameleyi ayrı tutmakta zorluklar olduğunda uygundur. Gerek duyulan toplam örneklem genişliği daha büyüktür, çünkü analiz gruplara dayandırılır (9, 27).

Doğru rasgelelik için verilen bu yöntemlerden hangisinin kullanılacağı deneme düzenine bağlıdır. Doğru rasgelelik yöntemleri hala, seçim yanlılığı ve etki karışımından sakınmak için tek yoldur.

Minimizasyon veya blok rasgelelik yöntemlerinde, bir araştırmacı sıradaki atamayı sezebildiğinden, basit rasgelelik dışındaki herhangi bir yöntemin kullanılmasındaki asıl problem atama gizlemenin bozulması olasılığına izin vermesidir. Atamanın gizlenmesi rasgeleliğin özündedir (27, 28). Herhangi bir denemede bir araştırmacı için sıradaki atamanın ne olacağını tahmin edebilmek imkansız olmalıdır. Eğer deneme yeterli büyüklükteyse, hiç bir işlem gerekli olmamalıdır.

2.3.9. Körleme

Rasgelelik kavramı, çalışmanın istatistiksel anlamlılık değerlendirmesindeki yanlılığı önlemek içindir, ancak muameleyle ilgili bilgilerden dolayı kayıt, değerlendirme, veri işleme ve istatistik analizdeki öznel kararlar nedeniyle ortaya çıkabilecek olan hiç bir yanlılık olmayacağını garanti etmez. Bu tip yanlılıklar doğrudan veya dolaylı olarak muameleyle ilgili olduğundan, muamele etkisi hakkındaki yorumu ciddi şekilde çarpıtabilir. Pratikte böyle bir yanlılığı ve onun muamelenin değerlendirilmesine olan etkisini nicel olarak değerlendirmek oldukça zordur. Bu nedenle klinik denemelerde, muamelelerle ilgili bilgiyi bloke ederek bunun gibi yanlılıkları elemek bir zorunluluktur. Bu yaklaşım körleme olarak adlandırılır. Körleme; denemeyle ilgili olan kişilerin, hastalara atanan muameleler hakkında herhangi bir bilgilerinin olmadığı deneysel bir durum olarak tanımlanır (2, 29). Körleme bazı kaynaklarda maskeleyme olarak da adlandırılmaktadır (21, 29).

Temel olarak klinik denemelerde kullanılan körleme dört tiptir (2, 7, 29):

1. Körleme Yok (Open Label)
2. Tek Kör (Single Blind)
3. Çift Kör (Double Blind)
4. Üçlü Kör (Triple Blind)

Bir körleme yok çalışması, hiçbir körlemenin kullanılmadığı bir klinik denemedir. Yani, hem araştırmacı hem de hasta, hastanın hangi muameleyi aldığı hakkında bilgi sahibidir (2, 29). Hastaların hangi muameleleri aldıklarının farkında olmaları, psikolojik olarak aldıkları muamele lehine tepki verebileceğinden, bu tip çalışmalarda ciddi bir yanlılık meydana gelecektir.

Tanım olarak tek kör çalışma, hasta veya araştırmacının hastanın atamasına kör olduğu tiptir (2, 29). Pratikte tek kör deneme, sadece hastanın kendi muamele atamasının farkında olmadığı bir deneme olarak ifade edilir. Körlemenin olmadığı denemelerle karşılaştırıldığında tek kör çalışmalar kesin bir kontrol derecesi ve klinik

denemelerin geçerliliği yönünden bir güvence verir. Bununla birlikte, arařtırmacı hastanın hangi muameleyi aldığını bildiğinden kendi klinik deęerlendirmesinde yanlılık ortaya çıkabilir.

Çift kör deneme, ne hastalar ne de arařtırmacının hastanın muamele atamasının farkında olmadığı bir denemedir. Hastalar ve arařtırmacıya ek olarak çalışmayla ilgili klinik proje takımının tüm üyeleri de körlenmişse bu klinik deneme üçlü kör olarak adlandırılır (2, 7, 29). Hastaların muamele atamasına ek olarak, körlük denemenin tüm sonuçlarının gizlenmesinde de uygulanır. Pratikte, proje üyelerinin genellikle hastanın verilerine erişimi olmasına rağmen, genellikle her bir hasta için muamele atamasının farkında değildirler. Körlük tipleri arasında üçlü kör denemeler, bir kontrollü klinik denemenin geçerliliği için en yüksek dereceyi sağlayabilir ve arařtırılan muamelenin etkinlik ve güvenlik deęerlendirmesi için en kesin ve yansız kanıtı sağlar.

2.4. KLİNİK DENEMELERDE DENEME DÜZENLERİ

2.4.1. Klinik Deneme Düzenleri

Uygun bir istatistiksel deneme düzeni seçmedeki ilk adım kurulacak klinik denemenin amacını belirlemektir. Bir klinik denemenin amacı, genellikle çalışılan tedaviye yönelik muameleyle ilgili bir veya daha fazla bilimsel veya tıbbi soruya cevap vermektir (2). Çalışmanın amacı tanımlandıktan sonra, amaçlanan klinik deneme için uygun bir deneme düzeni seçilebilir. Deneme düzeninin yanlış seçimi, işe yaramayan bir çalışmayla sonuçlanabildiğinden, iyi bir istatistiksel deneme düzeni zorunlu ilk şart olarak görülmektedir. Uygun bir deneme düzeni seçiminden önce sorulması zorunlu olan sorular, çalışmanın amacı, çalışılan muamelenin doğası, araştırma altındaki hastalığın durumu ve buna benzer konuları kapsamaktadır (2).

Çalışmanın amacı açık şekilde ifade edildikten sonra, çalışılacak konuları belirlemek önemlidir. Bu konular uygulama şekli, doz miktarı, dozaj şekli ve beklenen yanıt gibi konulardır (2). Bir yanıt araştırılırken, optimal tedavi etkisi için muamelenin uygun bir şekli ve miktarının uygulanması önemlidir. Ayrıca seçilen uygulama miktarı çalışılan muamelenin etkinlik ve güvenlik değerlendirmesini etkileyebilir. Düşük dozda bir uygulama daha iyi bir güvenlik profili gösterebilirken klinik anlamlılık etkisini meydana getiremeyebilir ve dolayısıyla etkinlik için yeterli olmayabilir. Yüksek dozda bir uygulama ise ciddi bir yan etkiye neden olabilir.

İdeal olarak, en uygun deneme düzeni, sadece tıbbi açıdan değil istatistiksel bakış açısı ve pazarlama gibi farklı açılardan da önemli olan konulara göre seçilmiş olanıdır (2). Pratikte böyle bir deneme düzeni mevcut olmayabilir. Çünkü farklı bakış açılarından önemli olan ve dikkate alınan konular, daima deneme düzeni seçiminde kısıtlamalara neden olacaktır. Bir klinik deneme planlanırken, uygun bir deneme düzeni seçilmeden önce, çalışma yapısına uygun olabilecek istatistiksel deneme düzenlerinin avantaj ve dezavantajlarının görece olarak karşılaştırılması gerekir (2, 11).

2.4.2. Klinik Denemelerin Sınıflandırılması

Temel olarak, fonksiyonlarına ve performans özelliklerine bağlı olarak birkaç farklı klinik deneme tipi vardır, fakat bunlar birbirinden tamamen ayrıdır. Bu farklı klinik deneme tipleri şu şekilde sınıflandırılabilir (2):

1. Çok Merkezli Denemeler (Multicenter Trials)
2. Daha Etkin Olma Denemeleri (Superiority Trials)
3. Ardışık Denemeler (Sequential Trials)
4. Aktif Kontrol ve Eşdeğerlik/Daha Kötü Olmama Denemeleri (Active Control and Equivalence/Noninferiority Trials)
5. Doz-Yanıt Denemeleri (Dose-Response Trials)
6. Kombinasyon Denemeleri (Combination Trials)
7. Köprüleme Çalışmaları (Bridging Studies)
8. Aşı Klinik Denemeleri (Vaccine Clinical Trials)

Çok merkezli denemeler klinik sonuçların çalışma merkezleri karşısında hedef hasta popülasyonuna genelleştirilebilirliğini ve kopyalanmasını sağlar. Eğer hedeflenen klinik denemenin amacı test muamelesinin birincil klinik bitim noktaları bakımından eş zamanlı bir kontrolden daha iyi olduğunu göstermekse, bu daha etkin olma denemeleri olarak ifade edilir. Çoğu vakalarda aktif kontrol denemeleri çalışmadaki denekler çok şiddetli hastalığa ya da yaşamı tehdit edici hastalıklara sahip olduğunda zorunlu olarak yeni bir muamelenin etkinliğini test etmek için yürütülmektedir. Başka bir ifadeyle aktif kontrol denemelerinin amaçlarından birisi de test muamelesinin etkinliğinin şimdiki aktif kontrolüne benzer olduğunu veya ondan daha aşağı olmadığını doğrulamaktır. Bu amacı taşıyan aktif kontrol denemeleri eşdeğerlik/daha kötü olmama denemeleri olarak ifade edilir. Daha kötü olmama denemeleri, aktif kontrolün daha kötü olmama etkinliğini göstererek test muamelesi ve aktif kontrol arasında tedavisel eşdeğerlik kurmak için düzenlenmektedir. Faz II klinik gelişim süresince, doz-yanıt denemeleri genellikle test muamelesinin tedavisel çerçevesini ve tedavisel çerçeve içindeki etkinlik/güvenlik ve dozlar arasındaki ilişkiyi belirlemek için kurulmaktadır. Çalışılan test muamelesi, örneğin bir ilaç birden fazla etken maddeden oluştuğunda, potansiyel

ilaç-ilaç interaksiyonunu hesaba da katarak muamele etkisini değerlendirmek için kombinasyon denemeleri gereklidir. Yeni bir bölge üzerine yabancı verileri değerlendirmek için yeni bölgede köprüleme çalışmaları yürütülmektedir. Çünkü aşı klinik denemeleri için hedef hasta popülasyonu genellikle normal sağlıklı bireylerdir, aşı denemeleri için ırklar, güvenlik, deneme düzeni ve örneklem genişliği gibi konuların dikkate alınması gerekmektedir (2).

2.5. ARDIŞIK DENEME DÜZENLERİ

2.5.1. Ardışık Deneme Düzenlerinin Tarihsel Gelişimi

Bu alandaki öncü çalışmalar arasında, şansa dayalı sistemler üzerine çalışan Huyghens, James Bernoulli, DeMoivre, Laplace ve diğerlerinin 17. ve 18. yüzyıl çalışmaları sayılabilir. Ardışık deneme düzenlerinin ilk resmi uygulamaları, 1920'lerin sonlarında imalat alanında istatistiksel kalite kontrolü ile başlamıştır. Shewhart (1931) işlem kontrolü için kontrol kartlarını geliştirmiştir. Dodge & Romig (1929) test edilebilen ve etkin veya kusurlu olarak sınıflandırılabilen bileşenler için iki aşamalı kabul örnekleme planını tanımlamışlardır (3, 30).

Modern ardışık analiz teorisi, 1943'ten bu yana silah üretimi ve gelişimi için endüstriyel danışma gruplarına katılan Abraham Wald (1947) ve George Barnard (1946) tarafından yapılan çalışmalardan ileri gelmektedir (3). Konu üzerindeki büyük etkiyi Wald (1947) tarafından yapılan “ardışık olasılık oran testi (AOOT) (sequential probability ratio test (SPRT))” ile ilgili çalışma oluşturmuştur (30). Zelen (1969), bir tıbbi çalışmaya giren deneklerin muamelelere atanması için bir pilot çalışmaya dayanan “play-the-winner” kuralını ortaya atmıştır ve bu kurala çeşitli modifikasyonlar önerilmiştir (3).

Armitage & Schneiderman (1958), Schneiderman (1961), Dunnett (1961) ve Roseberry & Gehan (1964) tarafından normal dağılım gösteren veriler için iki aşamalı veya üç aşamalı işlemler geliştirilmiştir. Armitage (1954, 1958b, 1975) ve Bross (1952, 1958) ardışık metotların tıbbi alanda, özellikle karşılaştırmalı klinik denemeler için kullanımına öncülük etmişlerdir (30). Başlangıçta bu planlar tamamen ardışıktı ve yaygın kabul görmemekteydi, bunun nedeni çalışma boyunca yapılan gözlemlerin sürekli olarak değerlendirilmesinin genellikle pratik olmamasıdır. Cutler ve ark. (1966)'nın makalesinden sonra basılan tartışmasında Shaw, “block sequential analysis” terimini kullanarak bir klinik deneme için grup ardışık metotların kullanımını önermiştir (3). Klinik denemeler için resmi grup ardışık metotlara kayma 1970lere kadar

olmamıştır. Elfring & Schultz (1973), iki muameleyi karşılaştırmak için olan yöntemlerini tanımlamak için “grup ardışık deneme düzeni” terimini özellikle kullandılar; McPherson (1974), az sayıda ara analizleri olan klinik deneme verilerini analiz etmek için kullanılabilir, Armitage'nin tekrarlanan anlamlılık testleri olan, McPherson & Rowe (1969) önerdi; Canner (1977) periyodik analizleri olan bir survival çalışmasında bir test istatistiğinin kritik değerlerini bulmak için Monte Carlo simülasyonunu kullandı (3). Bununla birlikte, grup ardışık metotlar için en büyük katkı, I. tip hata ve güç gereksinimlerine uygun grup ardışık deneme düzenlerinin uygulanması için açık bir kılavuz veren Pocock (1977)'tan geldi. Pocock aynı zamanda, normal dağılım gösteren veriler için kullanılan tekrarlanan anlamlılık testlerinin anlamlılık düzeylerinin diğer yanıtlar ve durumların çeşitliliği için güvenilir şekilde kullanılabilirliğini göstererek yaklaşımın çok yönlülüğünü de kanıtladı (3, 30).

Pocock'un makalesinden kısa süre sonra çıkan bir makalede, O'Brien & Fleming (1979), grup ardışık testlerin farklı bir sınıfını önerdi (3, 30). Bu testler ilk ara analizlerde durmak için ulaşılması zor, oldukça yüksek sınırlara sahiptir ve eğer son (final) analize kadar deneme devam ederse sabit örneklem testlerine yakın olan bir karara varmaktadır. Slud & Wei (1982) ve Lan & DeMets (1983) tarafından yazılan daha sonraki makaleler, klinik denemelerde yaygın olarak karşılaşılan, endüstriyel örnekleme uygulamalarında ise sık karşılaşılan bir problem olmayan, grup genişliklerinin eşit olmadığı grup ardışık deneme düzenlerinin kullanılabilirliğini gösterdiklerinden çok önemlidir. Pocock (1977), O'Brien & Fleming (1979) ve Lan & DeMets (1983) tarafından yazılan makaleler özellikle etkin olanlardır ve son zamanlardaki metodolojik araştırmalar ve klinik deneme düzenlerindeki uygulamaların temeli için başlama noktasını şekillendirmişlerdir (3, 30).

Etkinlik ve güvenlik verilerinin ara analizlerle değerlendirilmesini sağlayan istatistiksel yöntemlerin gerekliliği, 1980'lerde geniş ölçekli klinik denemelerin çoğalan sayısı ve bir denemenin gidişatını kontrol etmek için VGİK'lerinin artan kullanımı ve etkisiyle artmıştır. Şimdi neredeyse tüm geniş ölçekli uzun dönem klinik denemelerin protokolleri, biriken sonuçları istatistiksel olarak izleme işlemi için grup ardışık yöntemlerin kullanımını içermektedir (3). 1980'den sonra istatistiksel metodolojide,

daha çok uygulamaya yönelik ilerlemeler meydana gelmiştir. Grup ardışık metotların gelişmesi ve kabulü, biyoistatistiksel metotlardaki gelişmelerden ve ara analizler için gerekli olan veri yönetimi ve veri analizlerini kolaylaştıran bilgisayar teknolojisindeki hızlı ilerleyişten de yararlanmıştır.

Armitage, McPherson & Rowe (1969), eğer her bir ara analizde, sabit örneklem genişliği testi kritik değerleri kullanılırsa I. tip hata olasılığının başlangıçta belirlenen düzeyine göre fazlasıyla artacağını göstermişlerdir. Bu konuyla ilgili olarak Pocock (1977), O'Brien & Fleming (1979), Fleming, Harrington & O'Brien (1984) ve Wang & Tsiatis (1987) tarafından yapılan çalışmalar, grup ardışık deneme düzenlerinde genel I. tip hata olasılığını korumak için kritik değerleri ayarlamayla ilgili metotlardır. Bundan sonraki hesaplamalar daha çok istenen bir güce ulaşmak için grup ardışık testin gerek duyduğu örneklem genişliğinin belirlenmesiyle ilgili metotlardır (3, 30).

2.5.2. Neden Ardışık Deneme Düzenleri

Verilerin deneme süresince düzenli olarak biriktiği herhangi bir araştırmada, denemenin erken sonlandırılması veya deneme düzeninde bazı değişiklikler yapılması gibi gereklilikler ortaya çıkabilir. Bir deneme süresince toplanan verilerin ara analizlerinin yapılmasını gerektirecek nedenler temel olarak üç sınıfa ayrılabilir: “**etik**”, “**yönetimsel**” ve “**ekonomik**” nedenler (3, 20, 31, 32). Deneklerin insanlar olduğu klinik denemelerde, bireylerin güvensiz, etkin olmayan veya eşdeğer muamelelere maruz kalmalarını engellemek için sonuçların izlenmesi etik bir gerekliliktir. Ayrıca, test edilmek için bekleyen ümit verici başka muameleler de olduğunda bir denemeyi mümkün olduğu kadar çabuk sonlandırmak da etik bir zorunluluktur. Ara analizlerin yapılması için yönetimsel nedenler arasında, denemenin planlandığı şekilde yürütülüp yürütülmediği, deneklerin doğru popülasyondan olup olmadığı, deneklerin uygunluk kriterleriyle uyumlu olup olmadığı ve muamelelerin deneme protokolünde tanımlandığı şekilde uygulanıp uygulanmadığı gibi konuların kontrolü sayılabilir. Denemenin başında fark edilmeyen ya da fark edilemeyecek olan bazı problemlerin varlığının tespiti ve çok fazla masraf yapılmadan önce düzeltilebilmesi de yine ara analizlerle sonuçların erken

incelenmesi sayesinde olacaktır. Bunların yanında, ara analizlerle çalışma sonuçlarının erken incelenmesi, deneme düzeni seçilirken kabul edilen ön şartların da kontrol edilebilmesini sağlayacaktır.

Ardışık deneme düzenleri başlangıçta daha çok ekonomik yararlar elde etmek için geliştirilmiştir. Pozitif sonucu olan bir deneme için erken sonlandırma yeni bir ürünün daha çabuk kullanılabilceği anlamındadır. Eğer ara analizlerde bir pozitif sonuca ulaşırsa, denemenin erken durdurulması kaynakların boşa harcanmasını engelleyecektir. Ardışık deneme düzenleri standart sabit örneklem genişliği deneme düzenleri ile karşılaştırıldığında tipik olarak örneklem genişliği, zaman ve maliyet açısından tasarruf sağlayacaktır (3, 20, 31, 32).

Yüzlerce ya da binlerce denek içeren büyük denemelerde, her bir denek verilerinden sonra elde edilen sonuçları değerlendirmek pratik veya gerekli olmayabilir. Hasta sayısı daha az ve kayıt süresi daha kısa olan denemeler için, az aralıklı olarak her bir hastadan ya da hasta grubundan sonra izleme işleminin maliyeti ve uygulanması daha uygun olacaktır. Büyük veya küçük olsun bir denemenin erken aşamalarında, sonuçlar nadiren karar vermeye yetecek seviyede olur. Bu nedenle her bir hastadan ya da hasta grubundan sonra yapılan ilk ara analizler gereksiz gibi görünmektedir. Ancak bu erken analizler, beklenmeyen ya da fark edilemeyen bir riskin var olup olmadığını anlamak için bir tedbirdir. Bir deneme ilerledikçe ve veriler toplandıkça, muamele etkisinin ara analizlerle yapılan karşılaştırmaları daha güvenilir ve anlamlı olacaktır. Eğer yeni muamelenin daha etkin olduğu veya etkin olmadığı yönünde eğilimler ortaya çıkarsa, yapılması gereken şey daha yakın bir izleme yapılmasıdır. Bu özellikle etkin olmama yönündeki eğilimler için zorunludur (1, 19, 20, 33).

Bir denemeyi izlemek ve denemeyi durdurmaya veya devam etmeye karar vermek için birçok neden olabilmektedir. Tıbbi uygulamalarda, bunlar yan etkileri, mali giderleri, hastaların yaşam kalitesini veya denemenin başlangıcında bilinmeyen yeni ümit verici muamelelerin varlığını kapsar (3). En önemli neden ise, eğer muamelenin daha etkin olduğu veya etkin olmadığı yönünde sonuçlar ortaya çıkar ve bu eğilimleri gösteren test istatistikleri kritik değerlere yaklaşırsa, hastalara gereksiz yere etkin

olmayan bir muamele uygulanmaması için VGİK'nun bu eğilimleri değerlendirmek ve denemeyi sonlandırmak isteyebilmesidir (1). Deneyi durdurma kararı sadece istatistiksel değildir, daha doğrusu birçok öznel faktörleri kapsayan oldukça karmaşık bir karar işlemidir. Karmaşık ve öznel doğasından dolayı, bir denemeyi durdurmak veya devam etmek önemli bir tartışma nedeni olabilir. Genel sonuçlar çok güçlü ve inandırıcı, risk grupları çaprazlaması kadar birinci ve ikinci sonuçların çaprazlaması da tutarlı olmadıkça, eldeki sonuçlar denemeyi erken sonlandırmayı önermeye yetecek kadar güçlü olmayabilir. Verilerin değerlendirmesinde ve bir sonlandırma kararının verilmesinde çok dikkatli olunmalıdır. Bazı durumlarda, daha fazla veri ve sonuç olmasını sağlayacağından daha sık inceleme, bazı tutarsızlık problemlerini çözebilir (1).

2.5.3. Temel Ardışık Deneme Düzenleri

Ardışık deneme düzenleri, genel olarak tümüyle ardışık deneme düzenleri (fully sequential designs), grup ardışık deneme düzenleri (group sequential designs) ve esnek ardışık deneme düzenleri (flexible sequential designs) olarak sınıflandırılabilir (1).

2.5.3.1. Tümüyle Ardışık Deneme Düzenleri

Bu klasik ardışık metotların amacı, bir denemeye giren hastaların sayısını veya izlenme sürelerini, daha önce denemeye girmiş olanların sonuçlarından elde edilen karara dayanarak minimize etmektir. Klasik tümüyle ardışık deneme düzeninin gelecekteki gelişmelere temel oluşturan özelliği, her bir yeni hastanın ilk sonucu elde edildikten sonra muamelenin etkinliğinin değerlendirilmesidir. Burada, etki büyüklüğünü gösteren test istatistiğinin sınırları içinde, pozitif bir istatistik muamelenin daha etkin olduğunu ve negatif olanı da etkin olmadığını gösterir. Her bir yeni sonuçtan sonra, bir test istatistiği hesaplanır ve önceden belirlenen I. tip hata olasılığı düzeyinde, muameleler arasında farklılık olmadığı hipotezi, muamelelerden birinin daha etkin veya etkin olmayan bir etkiye sahip olduğu hipotezi ile karşılaştırılır. Ardışık deneme düzeni kritik değerini aşan büyük bir pozitif test istatistiği, muamelenin daha etkin olduğunu ve

istatistiksel olarak anlamlı olduğunu gösterecektir. Benzer şekilde, büyük bir negatif test istatistiği ise anlamlı bir etkin olmayan muamele etkisini gösterecektir. Arada kalan sonuçlar ise yeni hastaların eklenmesiyle denemenin devam etmesi gerektiğini ifade edecektir. Eğer deneme belirli bir süre geçmesine rağmen herhangi bir yönde eğilim göstermeyen test istatistikleriyle sonuçlanarak devam ederse, bazı ardışık deneme düzenleri muamele etkisi olmadığı kararıyla denemeyi sonlandırmayı önermektedir. Uygun deneme düzeni seçimi ve yeterli örneklem genişliği belirlenmesi şartları altında bu sonuç muamele eşdeğerliğinin kanıtı olarak dikkate alınmaktadır (1, 3).

2.5.3.2. Grup Ardışık Deneme Düzenleri

Her bir yeni hasta sonucundan sonra bir denemeyi değerlendirmek her zaman pratik veya gerekli olmayabilir. Bir denemenin maliyetleri açısından birden fazla sonuç veya sonuç çiftlerinin analizi için verileri güncellemek neredeyse imkansız olabilir. Ayrıca, her bir güncellemeden sonra bir komite toplantısı yapmanın maliyeti yüksek, uygulaması da zordur. Bu tip konuların bir sonucu olarak klasik tümüyle ardışık stratejinin, grup ardışık deneme düzeni olarak ifade edilen bir modifikasyonu ortaya çıkmaktadır. Grup ardışık deneme düzenleri, genişliği önceden belirlenen her bir yeni hasta grubundan sonra toplanan verilerin değerlendirilmesi temeline dayanmaktadır (1). Periyodik ara analizler yapılırken I. tip hata olasılığını kontrol eden istatistiksel kriterler geliştirilmiştir. Tümüyle ardışık deneme düzenine benzer şekilde, test istatistiği hesaplanır ve grup ardışık testin kritik değeriyle karşılaştırılır. Bu kritik değerler önceden belirlenen sayıda ara analizler ve I. tip hata olasılığı seçimlerine göre değişmektedir. Yaygın olarak kullanılan grup ardışık testler, Pocock ve O'Brien & Fleming tarafından önerilen testlerdir (31, 32).

Grup ardışık deneme düzenlerinin bir yararı, ara analizlerin periyodik olarak güncellenebilmesi ve inceleme için VGİK'na sunulabilmesidir. Grup ardışık deneme düzenleri ortalamaları, oranları veya hayatta kalma eğrilerini karşılaştıran yaygın test istatistikleri için geliştirilmiştir. Grup ardışık deneme düzenleri, ara analizlerin sayısının

önceden belirlenip sabit kalmasını ve denemeye giren deneklerin sayısı bakımından eşit aralıklarla yapılmasını gerektirmektedir (1).

2.5.3.3.Esnek Ardışık Deneme Düzenleri

Bir alfa kullanan fonksiyon (alpha spending function) yoluyla, grup ardışık deneme düzenlerinin esnek bir modifikasyonu sağlanmaktadır. İstatistiksel test işleminin alfa düzeyi, I. tip hata olasılığıdır. Alfa kullanan fonksiyon, denemenin sonunda, önceden belirlenen toplam I. tip hata olasılığı korunacak şekilde, her bir ara (interim) analizde dikkate alınacak olan I. tip hata olasılığını belirler. Bu ardışık metot, klasik ardışık deneme düzenleri ve grup ardışık deneme düzenleri için olduğu gibi toplam I. tip hata olasılığı düzeyini kontrol eder. Gerçekte, grup ardışık metotlar, bu yaklaşımın özel bir durumudur. Alfa kullanan fonksiyon seçimi deneme süresince değiştirilmemelidir. Alfa kullanan fonksiyon, ara analizlerin sonuçları muamelenin daha etkin olduğu ya da etkin olmadığı yönündeki istatistiksel kriterlere ulaşmaya yakın olduğunda, ara analizlerin sıklığının ve zamanının değişmesine olan esnekliği nedeniyle, grup ardışık deneme düzenlerinden daha esnektir. Esnek alfa kullanan fonksiyon (flexible alpha spending function) yaklaşımı, denk gelen alfa kullanan fonksiyon kullanıldığında yaklaşık olarak tümüyle ardışık deneme düzenindeki karar noktasına yaklaşacaktır. Alfa kullanan fonksiyon ardışık deneme düzeni, tümüyle ardışık deneme düzeni ve grup ardışık deneme düzeni arasında bir uzlaşmadır. Bu nedenlerden dolayı, alfa kullanan grup ardışık yaklaşımı şimdilerde yaygın olarak kullanılmaktadır.

2.6. GRUP ARDIŞIK DENEME DÜZENLERİ

2.6.1. İki Muameleyi Karşılaştırmak için Çift Taraflı Hipotez Testleri

Çift taraflı hipotez testlerinde, iki muamele arasında farklılık olmadığını ifade eden kontrol hipotezi (H_0), muameleler arasında farklılık olduğunu ifade eden ancak bu farklılığın ne yönde olduğu konusunda fikir vermeyen bir alternatif hipoteze (H_1) karşı test edilmektedir. Yani A ve B gibi iki muameleden elde edilen verilerin ortalamaları arasındaki farklılık karşılaştırılırken kurulacak hipotez takımı;

$$H_0 : \mu_A = \mu_B \text{ veya } H_0 : \mu_A - \mu_B = 0 \quad (2.1)$$

$$H_1 : \mu_A \neq \mu_B \text{ veya } H_1 : \mu_A - \mu_B \neq 0$$

şeklinde olacaktır (34). Muamele ortalamaları, varyansı bilinen bir normal dağılım gösterdiğinde kullanılacak test istatistiği Z istatistiğidir. Sabit örneklem testlerinde, elde edilen Z değeri “**kritik değer**” olarak adlandırılan belirli bir c değerinden büyük olduğunda H_0 hipotezi reddedilirken küçük olduğunda ise kabul edilmektedir. Burada kritik değer olan c değerinin belirlenmesinde, H_0 hipotezi doğrudurken ret kararı verilerek yapılan hata miktarı şeklinde tanımlanan “**I. tip hata**” (α) miktarı önemlidir. Bu hata miktarı genelde 0,05 olarak kabul edilirken özellikle ölüm riski veya geri dönüşümsüz zararları olan veya herhangi bir potansiyel riski olan klinik denemelerde ise 0,01 veya 0,001 olarak kabul edilmektedir. H_0 hipotezi yanlışken doğru karar verilerek H_0 hipotezinin reddedilmesi “**testin gücü**” olarak ifade edilir ve $(1 - \beta)$ ile gösterilir. Testin gücünü belirlemede önemli olan başlıca üç faktör; I. tip hata olasılığı (α), toplam örneklem genişliği (n) ve etki büyüklüğüdür (d). I. tip hata olasılığı çalışılacak konuya göre belirlenir. Etki büyüklüğü H_0 hipotezinde kabul edilen parametre değeri ile H_1 hipotezinde iddia edilen parametre değeri arasındaki fark olarak tanımlanabilir. Standardize edilmiş etki büyüklüğü ise bu farkın standart sapmaya bölünmesiyle elde edilen değerdir (16). Uygulamada genellikle belirli bir test gücünü

sağlayabilmek için kullanılması gereken minimum örneklem genişliği belirlenmektedir. Testin gücünün yüksek olması istenir ve genel olarak ulaşılması hedeflenen değerler 0,80 veya 0,90 güçtür. Burada β ise H_0 hipotezi yanlışken kabul edilerek yapılan hata miktarı olan “**II. tip hata**” miktarı olarak tanımlanır.

2.6.2. Sabit Örneklem Genişliği Testleri

İki muamele arasındaki farklılığı test etmek için yapılan ve sabit örneklem genişliği deneme düzeni kullanan bir çalışmada, A muamelesine atanan deneklerden elde edilen yanıtlar X_{Ai} ve B muamelesine atanan deneklerden elde edilen yanıtlar X_{Bi} ile gösterilsin. A muamelesindeki deneklerden elde edilen yanıtların $X_{Ai} \sim N(\mu_A, \sigma_A^2)$ şeklinde, B muamelesindeki deneklerden elde edilen yanıtların ise $X_{Bi} \sim N(\mu_B, \sigma_B^2)$ şeklinde ortalaması ve varyansı bilinen bir normal dağılım gösterdiği ve her bir muamelede n denek olduğu varsayımı altında iki muamele arasındaki farklılığı test etmek için kullanılacak Z istatistiği şu şekilde hesaplanır (34):

$$Z = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \quad (2.2)$$

Birbirinden bağımsız iki grubun karşılaştırılmasında her bir muamele grubunda gerek duyulan örneklem genişliği;

$$n_f = \frac{(Z_{\alpha/2} + Z_{\beta})^2 \cdot (\sigma_A^2 + \sigma_B^2)}{(\mu_A - \mu_B)^2} \quad (2.3)$$

formülüyle hesaplanır (3, 16, 17, 18, 35). Genel bir kural olarak buradan hesaplanacak değer uygulamada bir üst tamsayıya yuvarlanarak kullanılır.

2.6.3. Grup Ardışık Deneme Düzeni Testleri

Grup ardışık deneme düzenlerinde iki muamele karşılaştırılırken ilk olarak çalışılacak maksimum grup sayısı (K) ve gruptaki her bir muamele için gerek duyulan örneklem genişliği (m) belirlenir. Çalışmada yapılacak toplam analiz sayısı, $K - 1$ ara analiz ve 1 son analiz olmak üzere K tanedir. Denemeye alınacak maksimum denek sayısı ise $2mK$ kadardır. Denemede birden fazla analiz yapılacak olduğundan, toplam I. tip hata olasılığının (α) korunabilmesi oldukça önemli bir konudur ve denemenin başında belirlenen toplam I. tip hata olasılığının (α) ara analizlere dağıtılarak her bir ara analiz için kullanılacak I. tip hata olasılıklarının (α'_k) belirlenmesi gerekmektedir (1, 2, 3, 4, 5, 36). Denemenin başında seçilen I. tip hata olasılığını aşmamak için, I. tip hata düzeyi yapılacak analiz sayısına göre ayarlanarak her bir analizden sonra hesaplanan Z_k istatistiğinin karşılaştırılacağı kritik değer olan c_k değerleri belirlenir ($k = 1, 2, \dots, K$). Bu kritik değerler grafik üzerinde gösterildiğinde bir sınır oluşturur. $2m$ genişliğindeki her bir grup denemeye alındıktan sonra veriler analiz edilir ve her analiz için bir Z_k istatistiği hesaplanır. Hesaplanan Z_k değeri, o ara analiz için dikkate alınan kritik c_k değerinden büyük olursa iki muamele arasında farklılık olmadığını ifade eden kontrol hipotezi (H_0) reddedilerek deneme sonlandırılır, bu durum genellikle “**pozitif sonuç**” olarak ifade edilir. Aksi takdirde çalışmaya $2m$ genişliğinde bir grup daha alınarak denemeye devam edilir. Eğer H_0 hipotezini reddedecek bir sonuca ulaşılamaz ve deneme son gruba kadar devam ederse ve son analizde de yine $Z_K < c_K$ ise H_0 hipotezi kabul edilerek deneme sonlandırılır. Bu durum da genellikle “**negatif sonuç**” olarak ifade edilir (1, 3, 4, 5, 35).

Her bir analizde, o analiz için ardışık olmayan bir testten hesaplanan test istatistiği belirli bir kritik değerden büyük olursa H_0 hipotezi reddedileceği, aksi takdirde ise H_0 hipotezi kabul edileceğinden bu testler “**tekrarlanan anlamlılık testi (repeated significance test)**” olarak da ifade edilebilir (3, 20,35).

2.6.3.1. Pocock Testi

Pocock (1977), sabit bir anlamlılık düzeyinde tekrarlanan anlamlılık testi (repeated significance test) düşüncesini, çalışma süresince toplanan verilerin görel olarak daha az sayıda analiz edilmesi düşüncesine uyarlamıştır (31). Pocock testinde, deneklerin denemeye girişi, her bir muamelede m denek içerecek şekilde büyüklükleri eşit olarak ayarlanmış K gruba bölünür ve $2m$ genişliğindeki her bir grup çalışmaya alınıp gözlemler yapıldıktan sonra toplanan veriler analiz edilir. Her bir gözlem grubu çalışmaya alındıktan sonra test istatistiği şu şekilde hesaplanır (3, 35):

$$Z = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{m_A} + \frac{\sigma_B^2}{m_B}}}, \quad k = 1, 2, \dots, K \quad (2.4)$$

Pocock testinde, her bir ara analizden ve son analizden sonra hesaplanan test istatistiğinin karşılaştırılacağı kritik değer $C_p(K, \alpha)$ ile gösterilmektedir (3, 35). Burada $C_p(K, \alpha)$ kritik değerleri, toplam I. tip hata olasılığını (α) verecek şekilde seçilir. Pocock testinin formüsel gösterimi şu şekildedir (3, 35):

k . gruptan sonra ($k = 1, \dots, K - 1$),

- $|Z_k| \geq C_p(K, \alpha)$ ise H_0 hipotezi reddedilerek deneme sonlandırılır
- Aksi takdirde $k + 1$. grup ile denemeye devam edilir

K . gruptan sonra,

- $|Z_K| \geq C_p(K, \alpha)$ ise H_0 hipotezi reddedilerek deneme sonlandırılır
- Aksi takdirde H_0 hipotezi kabul edilerek deneme sonlandırılır

Buradaki $C_p(K, \alpha)$ kritik değerleri toplam I. tip hata olasılığı (α) ve yapılması planlanan toplam analiz sayısına (K) bağlı olarak değişmektedir ve deneme boyunca yapılacak her bir ara analiz ve son analiz için sabittir. Farklı α değerleri ve K sayıları

için $C_p(K, \alpha)$ değerleri Çizelge 2.1’de verilmiştir. Pocock testi için kullanılan kritik değerlerin oluşturdukları sınır şekline bir örnek olarak, $\alpha = 0,05$ I. tip hata olasılığı düzeyinde, 4 ara analiz ve 1 son analiz olmak üzere toplam $K = 5$ analiz içeren teste ait sınırlar Şekil 2.1’de gösterilmiştir.

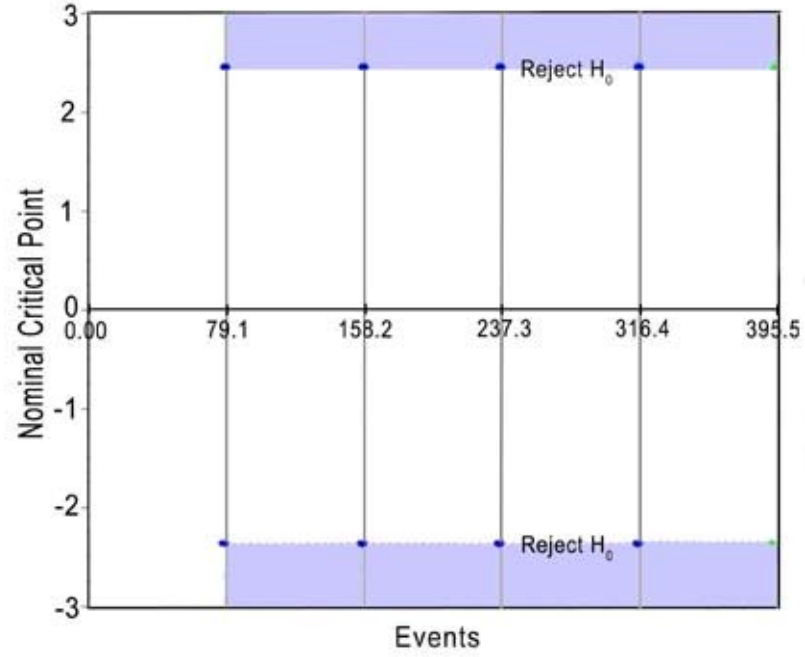
Çizelge 2.1 Pocock Testi $C_p(K, \alpha)$ Kritik Değerleri (3, 35)

K	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,10$
1	2,576	1,960	1,645
2	2,772	2,178	1,875
3	2,873	2,289	1,992
4	2,939	2,361	2,067
5	2,986	2,413	2,122
6	3,023	2,453	2,164
7	3,053	2,485	2,197
8	3,078	2,512	2,225
9	3,099	2,535	2,249
10	3,117	2,555	2,270
15	3,182	2,626	2,344
20	3,225	2,672	2,392

Pocock grup ardışık testinde her bir analiz grubunda çalışmaya alınacak muameleler için gerekli örneklem genişliğini belirlerken, grup ardışık testlerdeki maksimum örneklem genişliğinin sabit örneklem genişliği testindeki örneklem genişliğine oranı olan $R_p(K, \alpha, \beta)$ değerlerinin belirlenmesi gereklidir. Farklı K sayıları, α ve β değerleri için $R_p(K, \alpha, \beta)$ değerleri Çizelge 2.2’de verilmiştir.

Pocock testinde, K grup için gerek duyulan örneklem genişliği eşitlik (2.3)’teki n_f değerinin $R_p(K, \alpha, \beta)$ ile çarpılması sonucunda elde edilir. Her bir analiz grubunda her bir muameleye atanacak denek sayısı ise şu şekilde hesaplanır (3, 35).

$$m = (R_p(K, \alpha, \beta) \cdot n_f) / K \quad (2.5)$$



Şekil 2.1 Toplam 5 Analiz İçeren Pocock Testi Sınırları (5, 37)

Çizelge 2.2 Pocock Testi $R_p(K, \alpha, \beta)$ Değerleri (3, 35)

$(1-\beta)=0,80$				$(1-\beta)=0,90$			
K	$\alpha=0,01$	$\alpha=0,05$	$\alpha=0,10$	K	$\alpha=0,01$	$\alpha=0,05$	$\alpha=0,10$
1	1,000	1,000	1,000	1	1,000	1,000	1,000
2	1,092	1,110	1,121	2	1,084	1,100	1,110
3	1,137	1,166	1,184	3	1,125	1,151	1,166
4	1,166	1,202	1,224	4	1,152	1,183	1,202
5	1,187	1,229	1,254	5	1,170	1,207	1,228
6	1,203	1,249	1,277	6	1,185	1,225	1,249
7	1,216	1,265	1,296	7	1,197	1,239	1,266
8	1,226	1,279	1,311	8	1,206	1,252	1,280
9	1,236	1,291	1,325	9	1,215	1,262	1,292
10	1,243	1,301	1,337	10	1,222	1,271	1,302
15	1,272	1,338	1,381	15	1,248	1,305	1,341
20	1,291	1,363	1,411	20	1,264	1,327	1,367

2.6.3.2.O'Brien & Fleming Testi

O'Brien & Fleming (1979), sabit bir anlamlılık düzeyinde tekrarlanan anlamlılık testine bir alternatif olarak, deneme ilerledikçe her bir analizde H_0 hipotezini reddetmek için gerek duyulan anlamlılık düzeyinin arttığı bir test önermişlerdir (31, 32). Bu nedenle, bu grup ardışık test tipinde ilk analizlerde H_0 hipotezini reddetmek daha zorken sonraki analizlerde kolaylaşmaktadır (3, 19). Pocock testinde olduğu gibi deneklerin denemeye girişi, her bir muamelede m denek içerecek şekilde büyüklükleri eşit olarak ayarlanmış K gruba bölünür ve $2m$ genişliğindeki her bir grup çalışmaya alınıp gözlemler yapıldıktan sonra toplanan veriler analiz edilir. O'Brien & Fleming testi için test istatistiği de yine eşitlik (2.4)'teki gibi hesaplanır (3, 35).

O'Brien & Fleming testinde, son (final) analizden sonra hesaplanan test istatistiğinin karşılaştırılacağı kritik değer $C_B(K, \alpha)$ ile gösterilmektedir (3, 35). Herhangi bir ara analizde hesaplanan test istatistiğinin karşılaştırılacağı kritik değerler ise her bir ara analiz için farklıdır ve $C_B(K, \alpha)$ değerinin $\sqrt{K/k}$ ile çarpılmasıyla elde edilir. O'Brien & Fleming testinin formüsel gösterimi şu şekildedir (3, 35):

k . gruptan sonra ($k = 1, \dots, K - 1$),

- $|Z_k| \geq C_B(K, \alpha) \cdot \sqrt{K/k}$ ise H_0 hipotezi reddedilerek deneme sonlandırılır
- Aksi takdirde $k + 1$. grup ile denemeye devam edilir

K . gruptan sonra,

- $|Z_K| \geq C_B(K, \alpha)$ ise H_0 hipotezi reddedilerek deneme sonlandırılır
- Aksi takdirde H_0 hipotezi kabul edilerek deneme sonlandırılır

Burada da $C_B(K, \alpha)$ değeri, toplam I. tip hata olasılığı (α) ve yapılması planlanan toplam analiz sayısına (K) göre değişmektedir. Farklı α değerleri ve K sayıları için $C_B(K, \alpha)$ değerleri Çizelge 2.3'te verilmiştir. O'Brien & Fleming testi için kullanılan kritik değerlerin oluşturdukları sınır şekline bir örnek olarak, $\alpha = 0,05$ I. tip

hata olasılığı düzeyinde, 4 ara analiz ve 1 son analiz olmak üzere toplam $K = 5$ analiz içeren teste ait sınırlar Şekil 2.2’de gösterilmiştir.

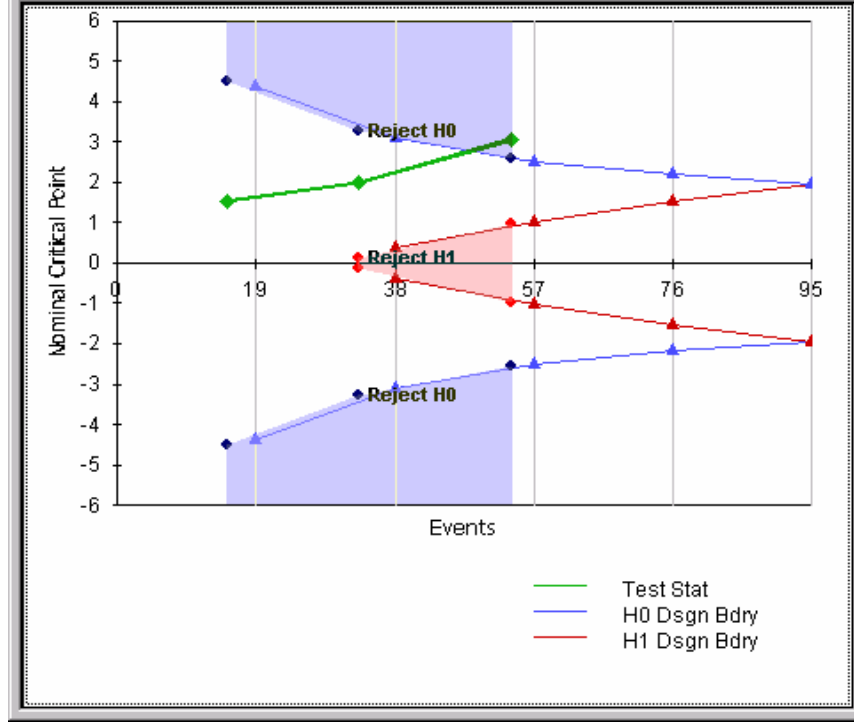
Çizelge 2.3 O’Brien & Fleming Testi $C_B(K, \alpha)$ Kritik Değerleri (3, 35)

K	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,10$
1	2,576	1,960	1,645
2	2,580	1,977	1,678
3	2,595	2,004	1,710
4	2,609	2,024	1,733
5	2,621	2,040	1,751
6	2,631	2,053	1,765
7	2,640	2,063	1,776
8	2,648	2,072	1,786
9	2,654	2,080	1,794
10	2,660	2,087	1,801
15	2,681	2,110	1,826
20	2,695	2,126	1,842

O’Brien & Fleming testinde her bir analiz grubunda çalışmaya alınacak muameleler için gerekli örneklem genişliğini belirlerken, Pocock testindeki $R_p(K, \alpha, \beta)$ değerlerine benzer şekilde $R_B(K, \alpha, \beta)$ değerleri belirlenir. Farklı K sayıları, α ve β değerleri için $R_B(K, \alpha, \beta)$ değerleri Çizelge 2.4’te verilmiştir.

K grup için her bir muamelede gerek duyulan örneklem genişliği belirlenirken yine eşitlik (2.3)’teki n_f değeri $R_B(K, \alpha, \beta)$ ile çarpılır. O’Brien & Fleming grup ardışık testinde, her bir analiz grubunda her bir muameleyle atanacak denek sayısı ise yine benzer şekilde hesaplanır (3, 35):

$$m = (R_B(K, \alpha, \beta) \cdot n_f) / K \quad (2.6)$$



Şekil 2.2 Toplam 5 Analiz İçeren O'Brien & Fleming Testi Sınırları (5, 38)

Çizelge 2.4 O'Brien & Fleming Testi $R_B(K, \alpha, \beta)$ Değerleri (3, 35)

K	$(1 - \beta) = 0,80$			$(1 - \beta) = 0,90$			
	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,10$	K	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,10$
1	1,000	1,000	1,000	1	1,000	1,000	1,000
2	1,001	1,008	1,016	2	1,001	1,007	1,014
3	1,007	1,017	1,027	3	1,006	1,016	1,025
4	1,011	1,024	1,035	4	1,010	1,022	1,032
5	1,015	1,028	1,040	5	1,014	1,026	1,037
6	1,017	1,032	1,044	6	1,016	1,030	1,041
7	1,019	1,035	1,047	7	1,018	1,032	1,044
8	1,021	1,037	1,049	8	1,020	1,034	1,046
9	1,022	1,038	1,051	9	1,021	1,036	1,048
10	1,024	1,040	1,053	10	1,022	1,037	1,049
15	1,028	1,045	1,058	15	1,026	1,042	1,054
20	1,030	1,047	1,061	20	1,029	1,045	1,057

2.6.3.3.Wang & Tsiatis Testi

Wang & Tsiatis (1987), diğer testlerden farklı olarak bir Δ parametresi ile özelleştirilen ve Pocock ve O'Brien & Fleming testlerinin kritik değerlerini bu parametrenin belirli değerleri için hesaplanan kritik değerlere eşit durumlar olarak kapsayan bir test ailesi önerdiler. Wang & Tsiatis testinde de yine deneklerin denemeye girişi, her bir muamelede m denek içerecek şekilde büyüklükleri eşit olarak ayarlanmış K gruba bölünür ve $2m$ genişliğindeki her bir grup çalışmaya alınıp gözlemler yapıldıktan sonra toplanan veriler analiz edilir. Test istatistiğinin hesaplanması da yine eşitlik (2.4)'teki gibidir (3, 35).

Wang & Tsiatis testinde, son analizden sonra hesaplanan test istatistiğinin karşılaştırılacağı kritik değer $C_{WT}(K, \alpha, \Delta)$ ile gösterilmektedir (3, 35). Herhangi bir ara analizde hesaplanan test istatistiğinin karşılaştırılacağı kritik değerler ise her bir ara analiz için farklıdır ve $C_{WT}(K, \alpha, \Delta)$ değerinin $(k/K)^{\Delta-1/2}$ ile çarpılmasıyla elde edilir. Testin formüsel gösterimi şu şekildedir (3, 35):

k . gruptan sonra ($k = 1, \dots, K - 1$),

- $|Z_k| \geq C_{WT}(K, \alpha, \Delta) \cdot (k/K)^{\Delta-1/2}$ ise H_0 hipotezi reddedilerek deneme sonlandırılır
- Aksi takdirde $k + 1$. grup ile denemeye devam edilir

K . gruptan sonra,

- $|Z_K| \geq C_{WT}(K, \alpha, \Delta)$ ise H_0 hipotezi reddedilerek deneme sonlandırılır
- Aksi takdirde H_0 hipotezi kabul edilerek deneme sonlandırılır

$\Delta = 0,50$ alındığında Wang & Tsiatis testinin sınırları Pocock testiyle, $\Delta = 0$ alındığında ise O'Brien & Fleming testiyle aynıdır. 0-0,5 arasındaki Δ değerleri ise arada kalan sınırlara sahip testleri vermektedir. $C_{WT}(K, \alpha, \Delta)$ değerleri de yine toplam I. tip hata olasılığı (α) ve yapılması planlanan toplam analiz sayısına (K) bağlı olarak

değişmektedir. $\alpha = 0,05$ I. tip hata olasılığı düzeyinde, farklı K sayıları ve Δ parametreleri için $C_{WT}(K, \alpha, \Delta)$ değerleri Çizelge 2.5'te verilmiştir.

Çizelge 2.5 Wang & Tsiatis Testi $C_{WT}(K, \alpha, \Delta)$ Kritik Değerleri (3, 35)

K	$\Delta = 0,10$	$\Delta = 0,25$	$\Delta = 0,40$
1	1,960	1,960	1,960
2	1,994	2,038	2,111
3	2,026	2,083	2,186
4	2,050	2,113	2,233
5	2,068	2,136	2,267
6	2,083	2,154	2,292
7	2,094	2,168	2,313
8	2,104	2,180	2,329
9	2,113	2,190	2,343
10	2,120	2,199	2,355
15	2,146	2,229	2,397
20	2,162	2,248	2,423

$\alpha = 0,05$ I. tip hata olasılığı düzeyinde, farklı K sayıları, β ve Δ değerleri için Çizelge 2.6'da verilen $R_{WT}(K, \alpha, \beta, \Delta)$ değerleri, gerekli örneklem genişliğinin belirlenmesinde Pocock ve O'Brien & Fleming testleri için verilen $R_p(K, \alpha, \beta)$ ve $R_b(K, \alpha, \beta)$ değerleriyle aynı şekilde kullanılır. Wang & Tsiatis testinde de K grup için her bir muamelede gerek duyulan örneklem genişliği, eşitlik (2.3)'teki n_f değerinin $R_{WT}(K, \alpha, \beta, \Delta)$ ile çarpılması sonucu elde edilir. Her bir analiz grubunda her bir muameleyle atanacak denek sayısı ise

$$m = (R_{WT}(K, \alpha, \beta, \Delta) \cdot n_f) / K \quad (2.7)$$

formülüyle hesaplanır (3, 35).

Çizelge 2.6 Wang & Tsiatis Testi $R_{WT}(K, \alpha, \beta, \Delta)$ Değerleri (3, 35)

$(1-\beta)=0,80$				$(1-\beta)=0,90$			
K	$\Delta=0,10$	$\Delta=0,25$	$\Delta=0,40$	K	$\Delta=0,10$	$\Delta=0,25$	$\Delta=0,40$
1	1,000	1,000	1,000	1	1,000	1,000	1,000
2	1,016	1,038	1,075	2	1,014	1,034	1,068
3	1,027	1,054	1,108	3	1,025	1,050	1,099
4	1,035	1,065	1,128	4	1,032	1,059	1,117
5	1,040	1,072	1,142	5	1,037	1,066	1,129
6	1,044	1,077	1,152	6	1,041	1,071	1,138
7	1,047	1,081	1,159	7	1,044	1,075	1,145
8	1,050	1,084	1,165	8	1,046	1,078	1,151
9	1,052	1,087	1,170	9	1,048	1,081	1,155
10	1,054	1,089	1,175	10	1,050	1,083	1,159
15	1,059	1,097	1,189	15	1,055	1,090	1,172
20	1,062	1,101	1,197	20	1,058	1,094	1,180

2.6.3.4. Haybittle-Peto Testi

Haybittle (1971) ve Peto ve ark. (1976), $k < K$ olan analizlerde, yani tüm ara analizlerde sadece $|Z_k| \geq 3$ olursa H_0 hipotezinin reddedildiği, basit bir ardışık izleme şekli önermişlerdir. Eğer ara analizlerin sayısı azsa H_0 hipotezini reddederek denemenin erken sonlandırılması olası değildir ve bu durum yaklaşık olarak ara analizlerin kullanılmadığı ve denemenin sonunda bir sabit örneklem genişliği testinin yapıldığı durumla aynıdır. Bununla birlikte, son analizde hesaplanan test istatistiğini karşılaştırmak için kullanılan kritik değer toplam I. tip hatayı verecek şekilde ayarlanabilir ve bu kabul edilebilecek bir yaklaşımdır. Diğer testlerle aynı şekilde, deneklerin denemeye girişi, her bir muamelede m denek içerecek şekilde büyüklükleri eşit olarak ayarlanmış K gruba bölünür ve $2m$ genişliğindeki her bir grup çalışmaya alınıp gözlemler yapıldıktan sonra toplanan veriler analiz edilir. Haybittle-Peto testinde de test istatistiği eşitlik (2.4)'teki gibi hesaplanır (3, 35).

Haybittle-Peto testinde, her bir ara analizden sonra hesaplanan test istatistiğinin karşılaştırılacağı kritik değer sabittir ($c_1 = \dots = c_{K-1} = 3$). Son analizden sonra hesaplanan

test istatistiğinin karşılaştırılacağı kritik değer ise $C_{HP}(K, \alpha)$ ile gösterilmektedir (3, 35). Testin formüsel gösterimi şu şekildedir (3, 35):

k . gruptan sonra ($k = 1, \dots, K - 1$),

- $|Z_k| \geq 3$ ise H_0 hipotezi reddedilerek deneme sonlandırılır
- Aksi takdirde $k + 1$. grup ile devam edilir

K . gruptan sonra,

- $|Z_K| \geq C_{HP}(K, \alpha)$ ise H_0 hipotezi reddedilerek deneme sonlandırılır
- Aksi takdirde H_0 hipotezi kabul edilerek deneme sonlandırılır

$c_k = 3$ seçiminin en uygun olduğu $\alpha = 0,05$ durumu incelendiğinde, hesaplamalar toplam analiz sayısı $K \geq 7$ olan bir deneme için, altıncı veya daha önceki bir ara analizde H_0 hipotezinin reddedilme olasılığının 0,0107 olduğunu göstermektedir (3). Yani $\alpha = 0,01$ düzeyinde genel I. tip hatayı başarmak mümkün değildir. $\alpha = 0,05$ I. tip hata olasılığı düzeyinde, farklı K sayıları için $C_{HP}(K, \alpha)$ değerleri Çizelge 2.7’de verilmiştir.

Haybittle-Peto testinde de her bir analiz grubunda çalışmaya alınacak muameleler için gerekli örneklem genişliğini belirlerken diğer test tiplerinde olduğu gibi, $R_{HP}(K, \alpha, \beta)$ değerleri belirlenir. $\alpha = 0,05$ I. tip hata olasılığı düzeyinde, farklı K sayıları ve β değerleri için $R_{HP}(K, \alpha, \beta)$ değerleri Çizelge 2.8’de verilmektedir. Yine, K grup için her bir muamelede gerek duyulan örneklem genişliği eşitlik (2.3)’teki n_f değerinin $R_{HP}(K, \alpha, \beta)$ ile çarpılmasıyla elde edilir. Her bir analiz grubunda her bir muameleyle atanacak denek sayısı yine aynı şekilde hesaplanır (3, 35):

$$m = (R_{HP}(K, \alpha, \beta) \cdot n_f) / K \quad (2.6.8)$$

Çizelge 2.7 Haybittle-Peto Testi $C_{HP}(K, \alpha)$ Kritik Değerleri (3, 35)

K	$C_{HP}(K, \alpha)$
1	1,960
2	1,967
3	1,975
4	1,983
5	1,990
6	1,997
7	2,003
8	2,010
9	2,016
10	2,021
15	2,046
20	2,068

Çizelge 2.8 Haybittle-Peto Testi $R_{HP}(K, \alpha, \beta)$ Değerleri (3, 35)

K	$(1 - \beta) = 0,80$	$(1 - \beta) = 0,90$
1	1,000	1,000
2	1,003	1,003
3	1,007	1,007
4	1,011	1,010
5	1,015	1,014
6	1,019	1,017
7	1,023	1,021
8	1,027	1,024
9	1,030	1,027
10	1,033	1,030
15	1,048	1,043
20	1,061	1,055

3. GEREÇ VE YÖNTEM

Bu tezin amacı, sabit örneklem genişliği deneme düzenlerine bir alternatif olarak ortaya çıkan grup ardışık deneme düzenlerinde uygulanan farklı test tiplerini karşılaştırmalı olarak incelemektir. Test tiplerinin birbirine göre avantaj ve dezavantajları, aynı koşullar altında gerek duyulan maksimum örneklem genişliği, denemeyi erken sonlandırma potansiyeli ve istenilen test gücü ile daha küçük muamele farklılıklarının tespit edilmesi yönlerinden ele alınmıştır. Karşılaştırmalar, iki muamele ortalamasının karşılaştırıldığı çift taraflı hipotez testleri için yapılmıştır. Bu doğrultuda 2 farklı I. tip hata (α), 2 farklı test gücü ($1 - \beta$), 14 farklı analiz grubu sayısı (K), 15 farklı etki büyüklüğü (d) ve 12 farklı varyans (σ^2) seçimlerinden oluşan toplam 10080 farklı kombinasyon oluşturulmuştur. Bu kombinasyonların oluşturulmasında dikkate alınan değerler şu şekildedir:

$$\alpha = 0,05 \text{ ve } 0,01$$

$$(1 - \beta) = 0,90 \text{ ve } 0,80$$

$$K = 1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 15 \text{ ve } 20$$

$$d = 0,5; 1,0; 1,5; 2,0; 2,5; 3,0; 3,5; 4,0; 4,5; 5,0; 6,0; 7,0; 8,0; 9,0; 10,0$$

$$\sigma^2 = 1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 15 \text{ ve } 20$$

Oluşturulan her bir kombinasyon için test tiplerinin gerek duydukları maksimum örneklem genişlikleri ve H_0 hipotezini reddetmek için kullandıkları kritik değerler belirlenmiştir. Bu kombinasyonların her birinde grup ardışık deneme düzenlerinde kullanılan dört farklı test tipi, birbirleriyle ve sabit örneklem genişliği deneme düzeni

testiyle karşılaştırılmış ve aynı I. tip hata (α), test gücü ($1-\beta$) ve analiz grubu sayısı (K) koşulları altında karşılaştırılan iki muamele arasındaki bilinen bir varyans ile normal dağılım gösteren belirli bir farklılığı tespit etme konusunda test tiplerinin birbirlerine göre durumları incelenmiştir. Karşılaştırmalarda testlerin belirli bir I. tip hata (α) düzeyinde H_0 hipotezini reddetmek için kullandığı kritik değerler, belirli bir test gücünü ($1-\beta$) sağlayabilmek için gerek duydukları maksimum örneklem genişlikleri ve denemeyi erken sonlandırarak sabit örneklem genişliği deneme düzeni testlerine göre daha az denek kullanma başarıları incelenmiştir.

Dikkate alınan her bir kombinasyon için her bir test tipinde ara (interim) analizlerden sonra hesaplanan test istatistiklerinin karşılaştırılacağı kritik değerler, Çizelge 2.1, Çizelge 2.3, Çizelge 2.5 ve Çizelge 2.7’de verilen her bir test tipinde son (final) analizden sonra hesaplanan test istatistiğinin karşılaştırılacağı c_K kritik değerlerinin her bir test tipinin bu kritik değerleri ara (interim) analizlere dağıtmak için kullandığı formüle göre düzenlenmesiyle hesaplanmıştır.

Gerek duyulan maksimum örneklem genişliği hesaplamalarında, dikkate alınan her bir kombinasyon için öncelikle eşitlik (2.4) yardımıyla sabit örneklem genişliği testinde gerek duyulan örneklem genişlikleri belirlenmiş, daha sonra bu eşitlikten hesaplanan değer, aynı I. tip hata (α) düzeyi ve test gücü ($1-\beta$) değerleri için, Çizelge 2.2, Çizelge 2.4, Çizelge 2.6 ve Çizelge 2.8’de verilen her bir testin kendi R değeriyle çarpılarak grup ardışık test tipleri için maksimum örneklem genişlikleri elde edilmiştir. Bölüm 2.6.3.4’te açıklandığı gibi $\alpha = 0,05$ düzeyinden daha düşük bir I. tip hata olasılığı sağlanamadığından Haybittle-Peto ve Wang & Tsatis testleri için maksimum örneklem genişlikleri sadece $\alpha = 0,05$ I. tip hata düzeyinde hesaplanabilmiştir. Bu nedenle de bu iki test tipi sadece $\alpha = 0,05$ I. tip hata olasılığı ile oluşturulan kombinasyonlarda karşılaştırmaya alınmış, bunun dışında kalan kombinasyonlarda Pocock ve O’Brien & Fleming testleri birbirleriyle ve sabit örneklem genişliği deneme düzeni testiyle karşılaştırılmıştır.

Her bir test tipinin, ara (interim) analizlerden sonra hesaplanan test istatistiklerinin karşılaştırılması için kullandığı kritik değerler hesaplanırken, dikkate alınan analiz grubu sayılarına (K) göre ortaya çıkan değişim söz konusu kritik değerlerin son analize doğru giderek azalması şeklinde olmaktadır. Bu durum tüm test tipleri ve tüm analiz grubu sayısı (K) değerleri için aynı şekilde ortaya çıkmaktadır. Dolayısıyla hesaplamalarda dikkate alınan analiz grubu sayısı (K) değerlerinin dışında kalan farklı analiz grubu sayıları (K) için ortaya çıkacak kritik değerler iterasyonla da bulunabilmektedir.

Gerek duyulan maksimum örneklem genişliği belirlenirken, artan I. tip hata düzeyi (α) ve etki büyüklüğü (d) için gerek duyulan örneklem genişliği de azalmakta, artan test gücü ($1-\beta$) ve varyans (σ^2) değerleri için ise gerek duyulan örneklem genişliği artmaktadır. Dolayısıyla çalışmanın başında oluşturulan 10080 farklı kombinasyonda, özellikle etki büyüklüğünün artması gerek duyulan örneklem genişliğinin oldukça azalmasına neden olmaktadır. $d = 3,0$ değerinden daha büyük etki büyüklüğü değerleri için gerek duyulan örneklem genişliği değerleri giderek azalmakta ve analiz grubu sayısı da dikkate alınınca her bir analiz grubunda gerek duyulan örneklem genişliği değerleri 1 sayısından daha küçük değerler alabilmektedir. $d = 5,0$ değerinden daha büyük etki büyüklüğü değerleri için ise gerek duyulan örneklem genişliği değerleri, henüz analiz grubu sayısına bölünmeden önce bile 1 sayısından daha küçük değerler alabilmektedir. Buna ek olarak hesaplamalarda, azalan varyans (σ^2) değerlerine de bağlı olarak gerek duyulan örneklem genişliği değerleri çok daha küçük sayılar olarak hesaplanabilmektedir. Teorik düzeyde bu hesaplamalar mümkün olmakta birlikte uygulamada bu sayılarla çalışabilmek mümkün değildir.

Sayılan tüm bu nedenlerle ve oluşturulan kombinasyonlara ilişkin bulgular oldukça fazla olduğundan, yapılan karşılaştırmaların sonuçları özetlenerek 2 farklı I. tip hata (α), 2 farklı test gücü ($1-\beta$), 5 farklı analiz grubu sayısı (K), 6 farklı etki büyüklüğü (d) ve 9 farklı varyans (σ^2) seçimlerinden oluşan kombinasyonlara ilişkin karşılaştırma sonuçları çizelge ve grafiklerle özetlenmiştir. Bu kombinasyonları içeren değerler şu şekildedir:

$$\alpha = 0,05 \text{ ve } 0,01$$

$$(1 - \beta) = 0,90 \text{ ve } 0,80$$

$$K = 1; 5; 10; 15 \text{ ve } 20$$

$$d = 0,5; 1,0; 1,5; 2,0; 2,5 \text{ ve } 3,0$$

$$\sigma^2 = 1; 2; 3; 4; 5; 6; 7; 8 \text{ ve } 9$$

Her bir test tipinin belirli bir I. tip hata (α) düzeyinde, farklı analiz grubu sayıları (K) için dikkate aldığı kritik değerler ve belirli bir test gücünü ($1 - \beta$) sağlayabilmek için gerek duydukları maksimum örneklem genişliklerine ilişkin karşılaştırma sonuçları çizelge ve grafiklerle özetlenmiştir. Grafik çizimlerinde Microsoft Office Standard Edition for Students and Teachers, Microsoft ® Office Excel 2003 modülünden yararlanılmıştır (39).

4. BULGULAR

Oluşturulan her bir kombinasyon için, sabit örneklem genişliği deneme düzeni testinin ve grup ardışık deneme düzenlerinde kullanılan dört farklı test tipinin her birinin belirli bir I. tip hata (α) düzeyinde, farklı analiz grubu sayıları (K) için H_0 hipotezini reddetmek için kullandıkları kritik değerler ve belirli bir test gücünü ($1-\beta$) sağlayabilmek için gerek duydukları maksimum örneklem genişliklerine ilişkin olarak elde edilen sonuçlar çizelge ve grafiklerle özetlenmiştir.

Çizelge 4.1'den Çizelge 4.8'e kadar olan ilk sekiz çizelgede dört farklı grup ardışık test tipi için belirli bir I. tip hata (α) düzeyinde, farklı analiz grubu sayıları (K) için H_0 hipotezini reddetmek için kullandıkları kritik değerler verilmiş, Şekil 5.1'den Şekil 5.8'e kadar olan ilk sekiz şekilde ise bu değerler grafikte gösterilmiştir. Çizelge 4.9'dan Çizelge 4.32'ye kadar olan sonraki yirmi dört çizelgede ise sabit örneklem genişliği testinin ve dört farklı grup ardışık test tipinin belirli bir I. tip hata (α) düzeyinde, farklı analiz grubu sayıları (K) ve belirli bir test gücü ($1-\beta$) için gerek duydukları maksimum örneklem genişlikler verilmiş, Şekil 5.9'dan Şekil 5.32'ye kadar olan yirmi dört şekilde ise bu değerler grafikte gösterilmiştir.

Çizelge 4.1 $\alpha = 0,05$ ve $K = 5$ için Kritik Değerler

K	C					
	C_P	C_B	$C_{WT},$ $\Delta = 0,10$	$C_{WT},$ $\Delta = 0,25$	$C_{WT},$ $\Delta = 0,40$	C_{HP}
1	2,413	4,562	3,937	3,194	2,663	3,000
2	2,413	3,226	2,984	2,686	2,485	3,000
3	2,413	2,634	2,357	2,427	2,386	3,000
4	2,413	2,281	2,261	2,259	2,318	3,000
5	2,413	2,040	2,068	2,136	2,267	1,990

Çizelge 4.2 $\alpha = 0,05$ ve $K = 10$ için Kritik Değerler

K	C					
	C_P	C_B	$C_{WT},$ $\Delta = 0,10$	$C_{WT},$ $\Delta = 0,25$	$C_{WT},$ $\Delta = 0,40$	C_{HP}
1	2,555	6,600	5,325	3,910	2,965	3,000
2	2,555	4,667	4,036	3,288	2,766	3,000
3	2,555	3,810	3,432	2,971	2,656	3,000
4	2,555	3,300	3,059	2,765	2,581	3,000
5	2,555	2,951	2,797	2,615	2,524	3,000
6	2,555	2,694	2,601	2,499	2,478	3,000
7	2,555	2,494	2,445	2,404	2,441	3,000
8	2,555	2,333	2,318	2,325	2,408	3,000
9	2,555	2,200	2,211	2,258	2,380	3,000
10	2,555	2,087	2,120	2,199	2,355	2,021

Çizelge 4.3 $\alpha = 0,05$ ve $K = 15$ için Kritik Değerler

K	C					
	C_P	C_B	$C_{WT},$ $\Delta = 0,10$	$C_{WT},$ $\Delta = 0,25$	$C_{WT},$ $\Delta = 0,40$	C_{HP}
1	2,626	8,172	6,340	4,387	3,143	3,000
2	2,626	5,778	4,805	3,689	2,932	3,000
3	2,626	4,718	4,085	3,333	2,816	3,000
4	2,626	4,086	3,641	3,102	2,736	3,000
5	2,626	3,655	3,330	2,934	2,675	3,000
6	2,626	3,336	3,096	2,803	2,627	3,000
7	2,626	3,089	2,911	2,697	2,587	3,000
8	2,626	2,889	2,760	2,608	2,553	3,000
9	2,626	2,724	2,633	2,533	2,523	3,000
10	2,626	2,584	2,524	2,467	2,496	3,000
11	2,626	2,464	2,429	2,409	2,473	3,000
12	2,626	2,359	2,346	2,357	2,451	3,000
13	2,626	2,267	2,272	2,310	2,432	3,000
14	2,626	2,184	2,206	2,268	2,414	3,000
15	2,626	2,110	2,146	2,229	2,397	2,046

Çizelge 4.4 $\alpha = 0,05$ ve $K = 20$ için Kritik Değerler

K	C					
	C_P	C_B	$C_{WT},$ $\Delta = 0,10$	$C_{WT},$ $\Delta = 0,25$	$C_{WT},$ $\Delta = 0,40$	C_{HP}
1	2,672	9,508	7,166	4,754	3,269	3,000
2	2,672	6,723	5,431	3,998	3,050	3,000
3	2,672	5,489	4,618	3,612	2,929	3,000
4	2,672	4,754	4,116	3,362	2,846	3,000
5	2,672	4,252	3,764	3,179	2,783	3,000
6	2,672	3,882	3,500	3,038	2,733	3,000
7	2,672	3,594	3,290	2,923	2,691	3,000
8	2,672	3,362	3,119	2,827	2,656	3,000
9	2,672	3,169	2,976	2,745	2,624	3,000
10	2,672	3,007	2,853	2,673	2,597	3,000
11	2,672	2,867	2,746	2,610	2,572	3,000
12	2,672	2,745	2,652	2,554	2,550	3,000
13	2,672	2,637	2,569	2,504	2,530	3,000
14	2,672	2,541	2,494	2,458	2,511	3,000
15	2,672	2,455	2,426	2,416	2,494	3,000
16	2,672	2,377	2,364	2,377	2,478	3,000
17	2,672	2,306	2,307	2,341	2,463	3,000
18	2,672	2,241	2,255	2,308	2,449	3,000
19	2,672	2,181	2,207	2,277	2,435	3,000
20	2,672	2,126	2,162	2,248	2,423	2,068

Çizelge 4.5 $\alpha = 0,01$ ve $K = 5$ için Kritik Değerler

K	C	
	C_P	C_B
1	2,986	5,861
2	2,986	4,144
3	2,986	3,384
4	2,986	2,930
5	2,986	2,621

Çizelge 4.6 $\alpha = 0,01$ ve $K = 10$ için Kritik Değerler

K	C	
	C_P	C_B
1	3,117	8,412
2	3,117	5,948
3	3,117	4,856
4	3,117	4,206
5	3,117	3,762
6	3,117	3,434
7	3,117	3,179
8	3,117	2,974
9	3,117	2,804
10	3,117	2,660

Çizelge 4.7 $\alpha = 0,01$ ve $K = 15$ için Kritik Değerler

K	C	
	C_P	C_B
1	3,182	10,383
2	3,182	7,342
3	3,182	5,995
4	3,182	5,192
5	3,182	4,644
6	3,182	4,239
7	3,182	3,925
8	3,182	3,671
9	3,182	3,461
10	3,182	3,284
11	3,182	3,131
12	3,182	2,997
13	3,182	2,880
14	3,182	2,775
15	3,182	2,681

Çizelge 4.8 $\alpha = 0,01$ ve $K = 20$ için Kritik Değerler

K	C	
	C_P	C_B
1	3,225	12,052
2	3,225	8,522
3	3,225	6,958
4	3,225	6,026
5	3,225	5,390
6	3,225	4,920
7	3,225	4,555
8	3,225	4,261
9	3,225	4,017
10	3,225	3,811
11	3,225	3,634
12	3,225	3,479
13	3,225	3,343
14	3,225	3,221
15	3,225	3,112
16	3,225	3,013
17	3,225	2,923
18	3,225	2,841
19	3,225	2,765
20	3,225	2,695

Çizelge 4.9 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 0,5$ için Örneklem Genişlikleri

		$d = 0,5$									
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$	
1	n_f	84,0	168,0	251,9	335,9	419,9	503,9	587,9	671,8	755,8	
	n_P	101,4	202,7	304,1	405,5	506,8	608,2	709,6	810,9	912,3	
	n_B	86,2	172,3	258,5	344,7	430,8	517,0	603,2	689,3	775,5	
	5	$n_{WT}, \Delta = 0,10$	87,1	174,2	261,3	348,4	435,4	522,5	609,6	696,7	783,8
		$n_{WT}, \Delta = 0,25$	89,5	179,0	268,6	358,1	447,6	537,1	626,7	716,2	805,7
		$n_{WT}, \Delta = 0,40$	94,8	189,6	284,4	379,3	474,1	568,9	663,7	758,5	853,3
	n_{HP}	85,2	170,3	255,5	340,6	425,8	510,9	596,1	681,3	766,4	
10	n_P	106,7	213,5	320,2	427,0	533,7	640,4	747,2	853,9	960,7	
	n_B	87,1	174,2	261,3	348,4	435,4	522,5	609,6	696,7	783,8	
	10	$n_{WT}, \Delta = 0,10$	88,2	176,4	264,5	352,7	440,9	529,1	617,3	705,4	793,6
		$n_{WT}, \Delta = 0,25$	91,0	181,9	272,9	363,8	454,8	545,7	636,7	727,6	818,6
		$n_{WT}, \Delta = 0,40$	97,3	194,7	292,0	389,3	486,7	584,0	681,3	778,7	876,0
		n_{HP}	86,5	173,0	259,5	346,0	432,5	519,0	605,5	692,0	778,5
15	n_P	109,6	219,2	328,8	438,4	548,0	657,6	767,2	876,8	986,4	
	n_B	87,5	175,0	262,5	350,0	437,5	525,0	612,6	700,1	787,6	
	15	$n_{WT}, \Delta = 0,10$	88,6	177,2	265,8	354,4	443,0	531,6	620,2	708,8	797,4
		$n_{WT}, \Delta = 0,25$	91,5	183,1	274,6	366,2	457,7	549,2	640,8	732,3	823,9
		$n_{WT}, \Delta = 0,40$	98,4	196,9	295,3	393,7	492,1	590,6	689,0	787,4	885,8
		n_{HP}	87,6	175,2	262,8	350,4	438,0	525,6	613,1	700,7	788,3
20	n_P	111,4	222,9	334,3	445,8	557,2	668,7	780,1	891,5	1003,0	
	n_B	87,8	175,5	263,3	351,0	438,8	526,6	614,3	702,1	789,8	
	20	$n_{WT}, \Delta = 0,10$	88,9	177,7	266,6	355,4	444,3	533,1	622,0	710,8	799,7
		$n_{WT}, \Delta = 0,25$	91,9	183,7	275,6	367,5	459,4	551,2	643,1	735,0	826,9
		$n_{WT}, \Delta = 0,40$	99,1	198,2	297,3	396,4	495,5	594,6	693,7	792,8	891,9
		n_{HP}	88,6	177,2	265,8	354,4	443,0	531,6	620,2	708,8	797,4

Çizelge 4.10 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 1,0$ için Örneklem Genişlikleri

		$d = 1,0$									
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$	
1	n_f	21,0	42,0	63,0	84,0	105,0	126,0	147,0	168,0	189,0	
	n_P	25,3	50,7	76,0	101,4	126,7	152,0	177,4	202,7	228,1	
	n_B	21,5	43,1	64,6	86,2	107,7	129,2	150,8	172,3	193,9	
	5	$n_{WT}, \Delta = 0,10$	21,8	43,5	65,3	87,1	108,9	130,6	152,4	174,2	195,9
		$n_{WT}, \Delta = 0,25$	22,4	44,8	67,1	89,5	111,9	134,3	156,7	179,0	201,4
		$n_{WT}, \Delta = 0,40$	23,7	47,4	71,1	94,8	118,5	142,2	165,9	189,6	213,3
	n_{HP}	21,3	42,6	63,9	85,2	106,4	127,7	149,0	170,3	191,6	
10	n_P	26,7	53,4	80,1	106,7	133,4	160,1	186,8	213,5	240,2	
	n_B	21,8	43,5	65,3	87,1	108,9	130,6	152,4	174,2	195,9	
	10	$n_{WT}, \Delta = 0,10$	22,0	44,1	66,1	88,2	110,2	132,3	154,3	176,4	198,4
		$n_{WT}, \Delta = 0,25$	22,7	45,5	68,2	91,0	113,7	136,4	159,2	181,9	204,6
		$n_{WT}, \Delta = 0,40$	24,3	48,7	73,0	97,3	121,7	146,0	170,3	194,7	219,0
		n_{HP}	21,6	43,3	64,9	86,5	108,1	129,8	151,4	173,0	194,6
15	n_P	27,4	54,8	82,2	109,6	137,0	164,4	191,8	219,2	246,6	
	n_B	21,9	43,8	65,6	87,5	109,4	131,3	153,1	175,0	196,9	
	15	$n_{WT}, \Delta = 0,10$	22,1	44,3	66,4	88,6	110,7	132,9	155,0	177,2	199,3
		$n_{WT}, \Delta = 0,25$	22,9	45,8	68,7	91,5	114,4	137,3	160,2	183,1	206,0
		$n_{WT}, \Delta = 0,40$	24,6	49,2	73,8	98,4	123,0	147,6	172,2	196,9	221,5
		n_{HP}	21,9	43,8	65,7	87,6	109,5	131,4	153,3	175,2	197,1
20	n_P	27,9	55,7	83,6	111,4	139,3	167,2	195,0	222,9	250,7	
	n_B	21,9	43,9	65,8	87,8	109,7	131,6	153,6	175,5	197,5	
	20	$n_{WT}, \Delta = 0,10$	22,2	44,4	66,6	88,9	111,1	133,3	155,5	177,7	199,9
		$n_{WT}, \Delta = 0,25$	23,0	45,9	68,9	91,9	114,8	137,8	160,8	183,7	206,7
		$n_{WT}, \Delta = 0,40$	24,8	49,5	74,3	99,1	123,9	148,6	173,4	198,2	223,0
		n_{HP}	22,1	44,3	66,4	88,6	110,7	132,9	155,0	177,2	199,3

Çizelge 4.11 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 1,5$ için Örneklem Genişlikleri

		$d = 1,5$									
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$	
1	n_f	9,3	18,7	28,0	37,3	46,7	56,0	65,3	74,6	84,0	
	n_P	11,3	22,5	33,8	45,1	56,3	67,6	78,8	90,1	101,4	
	n_B	9,6	19,1	28,7	38,3	47,9	57,4	67,0	76,6	86,2	
	5	$n_{WT}, \Delta = 0,10$	9,7	19,4	29,0	38,7	48,4	58,1	67,7	77,4	87,1
		$n_{WT}, \Delta = 0,25$	9,9	19,9	29,8	39,8	49,7	59,7	69,6	79,6	89,5
		$n_{WT}, \Delta = 0,40$	10,5	21,1	31,6	42,1	52,7	63,2	73,7	84,3	94,8
	n_{HP}	9,5	18,9	28,4	37,8	47,3	56,8	66,2	75,7	85,2	
10	n_P	11,9	23,7	35,6	47,4	59,3	71,2	83,0	94,9	106,7	
	n_B	9,7	19,4	29,0	38,7	48,4	58,1	67,7	77,4	87,1	
	10	$n_{WT}, \Delta = 0,10$	9,8	19,6	29,4	39,2	49,0	58,8	68,6	78,4	88,2
		$n_{WT}, \Delta = 0,25$	10,1	20,2	30,3	40,4	50,5	60,6	70,7	80,8	91,0
		$n_{WT}, \Delta = 0,40$	10,8	21,6	32,4	43,3	54,1	64,9	75,7	86,5	97,3
		n_{HP}	9,6	19,2	28,8	38,4	48,1	57,7	67,3	76,9	86,5
15	n_P	12,2	24,4	36,5	48,7	60,9	73,1	85,2	97,4	109,6	
	n_B	9,7	19,4	29,2	38,9	48,6	58,3	68,1	77,8	87,5	
	15	$n_{WT}, \Delta = 0,10$	9,8	19,7	29,5	39,4	49,2	59,1	68,9	78,8	88,6
		$n_{WT}, \Delta = 0,25$	10,2	20,3	30,5	40,7	50,9	61,0	71,2	81,4	91,5
		$n_{WT}, \Delta = 0,40$	10,9	21,9	32,8	43,7	54,7	65,6	76,6	87,5	98,4
		n_{HP}	9,7	19,5	29,2	38,9	48,7	58,4	68,1	77,9	87,6
20	n_P	12,4	24,8	37,1	49,5	61,9	74,3	86,7	99,1	111,4	
	n_B	9,8	19,5	29,3	39,0	48,8	58,5	68,3	78,0	87,8	
	20	$n_{WT}, \Delta = 0,10$	9,9	19,7	29,6	39,5	49,4	59,2	69,1	79,0	88,9
		$n_{WT}, \Delta = 0,25$	10,2	20,4	30,6	40,8	51,0	61,2	71,5	81,7	91,9
		$n_{WT}, \Delta = 0,40$	11,0	22,0	33,0	44,0	55,1	66,1	77,1	88,1	99,1
		n_{HP}	9,8	19,7	29,5	39,4	49,2	59,1	68,9	78,8	88,6

Çizelge 4.12 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 2,0$ için Örneklem Genişlikleri

		$d = 2,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	5,2	10,5	15,7	21,0	26,2	31,5	36,7	42,0	47,2
	n_P	6,3	12,7	19,0	25,3	31,7	38,0	44,3	50,7	57,0
	n_B	5,4	10,8	16,2	21,5	26,9	32,3	37,7	43,1	48,5
	$n_{WT}, \Delta = 0,10$	5,4	10,9	16,3	21,8	27,2	32,7	38,1	43,5	49,0
	$n_{WT}, \Delta = 0,25$	5,6	11,2	16,8	22,4	28,0	33,6	39,2	44,8	50,4
5	$n_{WT}, \Delta = 0,40$	5,9	11,9	17,8	23,7	29,6	35,6	41,5	47,4	53,3
	n_{HP}	5,3	10,6	16,0	21,3	26,6	31,9	37,3	42,6	47,9
	n_P	6,7	13,3	20,0	26,7	33,4	40,0	46,7	53,4	60,0
	n_B	5,4	10,9	16,3	21,8	27,2	32,7	38,1	43,5	49,0
	$n_{WT}, \Delta = 0,10$	5,5	11,0	16,5	22,0	27,6	33,1	38,6	44,1	49,6
10	$n_{WT}, \Delta = 0,25$	5,7	11,4	17,1	22,7	28,4	34,1	39,8	45,5	51,2
	$n_{WT}, \Delta = 0,40$	6,1	12,2	18,3	24,3	30,4	36,5	42,6	48,7	54,8
	n_{HP}	5,4	10,8	16,2	21,6	27,0	32,4	37,8	43,3	48,7
	n_P	6,8	13,7	20,5	27,4	34,2	41,1	47,9	54,8	61,6
	n_B	5,5	10,9	16,4	21,9	27,3	32,8	38,3	43,8	49,2
15	$n_{WT}, \Delta = 0,10$	5,5	11,1	16,6	22,1	27,7	33,2	38,8	44,3	49,8
	$n_{WT}, \Delta = 0,25$	5,7	11,4	17,2	22,9	28,6	34,3	40,0	45,8	51,5
	$n_{WT}, \Delta = 0,40$	6,2	12,3	18,5	24,6	30,8	36,9	43,1	49,2	55,4
	n_{HP}	5,5	10,9	16,4	21,9	27,4	32,8	38,3	43,8	49,3
	n_P	7,0	13,9	20,9	27,9	34,8	41,8	48,8	55,7	62,7
20	n_B	5,5	11,0	16,5	21,9	27,4	32,9	38,4	43,9	49,4
	$n_{WT}, \Delta = 0,10$	5,6	11,1	16,7	22,2	27,8	33,3	38,9	44,4	50,0
	$n_{WT}, \Delta = 0,25$	5,7	11,5	17,2	23,0	28,7	34,5	40,2	45,9	51,7
	$n_{WT}, \Delta = 0,40$	6,2	12,4	18,6	24,8	31,0	37,2	43,4	49,5	55,7
	n_{HP}	5,5	11,1	16,6	22,1	27,7	33,2	38,8	44,3	49,8

Çizelge 4.13 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 2,5$ için Örneklem Genişlikleri

		$d = 2,5$									
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$	
1	n_f	3,4	6,7	10,1	13,4	16,8	20,2	23,5	26,9	30,2	
	n_P	4,1	8,1	12,2	16,2	20,3	24,3	28,4	32,4	36,5	
	n_B	3,4	6,9	10,3	13,8	17,2	20,7	24,1	27,6	31,0	
	5	$n_{WT}, \Delta = 0,10$	3,5	7,0	10,5	13,9	17,4	20,9	24,4	27,9	31,4
		$n_{WT}, \Delta = 0,25$	3,6	7,2	10,7	14,3	17,9	21,5	25,1	28,6	32,2
		$n_{WT}, \Delta = 0,40$	3,8	7,6	11,4	15,2	19,0	22,8	26,5	30,3	34,1
	n_{HP}	3,4	6,8	10,2	13,6	17,0	20,4	23,8	27,3	30,7	
10	n_P	4,3	8,5	12,8	17,1	21,3	25,6	29,9	34,2	38,4	
	n_B	3,5	7,0	10,5	13,9	17,4	20,9	24,4	27,9	31,4	
	10	$n_{WT}, \Delta = 0,10$	3,5	7,1	10,6	14,1	17,6	21,2	24,7	28,2	31,7
		$n_{WT}, \Delta = 0,25$	3,6	7,3	10,9	14,6	18,2	21,8	25,5	29,1	32,7
		$n_{WT}, \Delta = 0,40$	3,9	7,8	11,7	15,6	19,5	23,4	27,3	31,1	35,0
		n_{HP}	3,5	6,9	10,4	13,8	17,3	20,8	24,2	27,7	31,1
15	n_P	4,4	8,8	13,2	17,5	21,9	26,3	30,7	35,1	39,5	
	n_B	3,5	7,0	10,5	14,0	17,5	21,0	24,5	28,0	31,5	
	15	$n_{WT}, \Delta = 0,10$	3,5	7,1	10,6	14,2	17,7	21,3	24,8	28,4	31,9
		$n_{WT}, \Delta = 0,25$	3,7	7,3	11,0	14,6	18,3	22,0	25,6	29,3	33,0
		$n_{WT}, \Delta = 0,40$	3,9	7,9	11,8	15,7	19,7	23,6	27,6	31,5	35,4
		n_{HP}	3,5	7,0	10,5	14,0	17,5	21,0	24,5	28,0	31,5
20	n_P	4,5	8,9	13,4	17,8	22,3	26,7	31,2	35,7	40,1	
	n_B	3,5	7,0	10,5	14,0	17,6	21,1	24,6	28,1	31,6	
	20	$n_{WT}, \Delta = 0,10$	3,6	7,1	10,7	14,2	17,8	21,3	24,9	28,4	32,0
		$n_{WT}, \Delta = 0,25$	3,7	7,3	11,0	14,7	18,4	22,0	25,7	29,4	33,1
		$n_{WT}, \Delta = 0,40$	4,0	7,9	11,9	15,9	19,8	23,8	27,7	31,7	35,7
		n_{HP}	3,5	7,1	10,6	14,2	17,7	21,3	24,8	28,4	31,9

Çizelge 4.14 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 3,0$ için Örneklem Genişlikleri

		$d = 3,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	2,3	4,7	7,0	9,3	11,7	14,0	16,3	18,7	21,0
	n_P	2,8	5,6	8,4	11,3	14,1	16,9	19,7	22,5	25,3
	n_B	2,4	4,8	7,2	9,6	12,0	14,4	16,8	19,1	21,5
	$n_{WT}, \Delta = 0,10$	2,4	4,8	7,3	9,7	12,1	14,5	16,9	19,4	21,8
	$n_{WT}, \Delta = 0,25$	2,5	5,0	7,5	9,9	12,4	14,9	17,4	19,9	22,4
5	$n_{WT}, \Delta = 0,40$	2,6	5,3	7,9	10,5	13,2	15,8	18,4	21,1	23,7
	n_{HP}	2,4	4,7	7,1	9,5	11,8	14,2	16,6	18,9	21,3
	n_P	3,0	5,9	8,9	11,9	14,8	17,8	20,8	23,7	26,7
	n_B	2,4	4,8	7,3	9,7	12,1	14,5	16,9	19,4	21,8
	$n_{WT}, \Delta = 0,10$	2,4	4,9	7,3	9,8	12,2	14,7	17,1	19,6	22,0
10	$n_{WT}, \Delta = 0,25$	2,5	5,1	7,6	10,1	12,6	15,2	17,7	20,2	22,7
	$n_{WT}, \Delta = 0,40$	2,7	5,4	8,1	10,8	13,5	16,2	18,9	21,6	24,3
	n_{HP}	2,4	4,8	7,2	9,6	12,0	14,4	16,8	19,2	21,6
	n_P	3,0	6,1	9,1	12,2	15,2	18,3	21,3	24,4	27,4
	n_B	2,4	4,9	7,3	9,7	12,2	14,6	17,0	19,4	21,9
15	$n_{WT}, \Delta = 0,10$	2,5	4,9	7,4	9,8	12,3	14,8	17,2	19,7	22,1
	$n_{WT}, \Delta = 0,25$	2,5	5,1	7,6	10,2	12,7	15,3	17,8	20,3	22,9
	$n_{WT}, \Delta = 0,40$	2,7	5,5	8,2	10,9	13,7	16,4	19,1	21,9	24,6
	n_{HP}	2,4	4,9	7,3	9,7	12,2	14,6	17,0	19,5	21,9
	n_P	3,1	6,2	9,3	12,4	15,5	18,6	21,7	24,8	27,9
20	n_B	2,4	4,9	7,3	9,8	12,2	14,6	17,1	19,5	21,9
	$n_{WT}, \Delta = 0,10$	2,5	4,9	7,4	9,9	12,3	14,8	17,3	19,7	22,2
	$n_{WT}, \Delta = 0,25$	2,6	5,1	7,7	10,2	12,8	15,3	17,9	20,4	23,0
	$n_{WT}, \Delta = 0,40$	2,8	5,5	8,3	11,0	13,8	16,5	19,3	22,0	24,8
	n_{HP}	2,5	4,9	7,4	9,8	12,3	14,8	17,2	19,7	22,1

Çizelge 4.15 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 0,5$ için Örneklem Genişlikleri

		$d = 0,5$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	62,7	125,4	188,2	250,9	313,6	376,3	439,0	501,8	564,5
	n_P	77,1	154,2	231,2	308,3	385,4	462,5	539,6	616,7	693,7
	n_B	64,5	129,0	193,4	257,9	322,4	386,9	451,3	515,8	580,3
	$n_{WT}, \Delta = 0,10$	65,2	130,5	195,7	260,9	326,1	391,4	456,6	521,8	587,1
	$n_{WT}, \Delta = 0,25$	67,2	134,5	201,7	268,9	336,2	403,4	470,7	537,9	605,1
5	$n_{WT}, \Delta = 0,40$	71,6	143,3	214,9	286,5	358,1	429,8	501,4	573,0	644,6
	n_{HP}	63,7	127,3	191,0	254,6	318,3	382,0	445,6	509,3	572,9
	n_P	81,6	163,2	244,8	326,4	408,0	489,6	571,2	652,8	734,4
	n_B	65,2	130,5	195,7	260,9	326,1	391,4	456,6	521,8	587,1
	$n_{WT}, \Delta = 0,10$	66,1	132,2	198,3	264,4	330,5	396,6	462,7	528,9	595,0
10	$n_{WT}, \Delta = 0,25$	68,3	136,6	204,9	273,2	341,5	409,8	478,1	546,4	614,7
	$n_{WT}, \Delta = 0,40$	73,7	147,4	221,1	294,8	368,5	442,2	515,9	589,6	663,3
	n_{HP}	64,8	129,6	194,4	259,2	323,9	388,7	453,5	518,3	583,1
	n_P	83,9	167,8	251,8	335,7	419,6	503,5	587,4	671,4	755,3
	n_B	65,5	131,1	196,6	262,2	327,7	393,3	458,8	524,3	589,9
15	$n_{WT}, \Delta = 0,10$	66,4	132,8	199,3	265,7	332,1	398,5	464,9	531,4	597,8
	$n_{WT}, \Delta = 0,25$	68,8	137,6	206,4	275,2	344,0	412,8	481,6	550,4	619,2
	$n_{WT}, \Delta = 0,40$	74,6	149,1	223,7	298,3	372,9	447,4	522,0	596,6	671,2
	n_{HP}	65,7	131,5	197,2	262,9	328,7	394,4	460,1	525,8	591,6
	n_P	85,5	171,0	256,5	341,9	427,4	512,9	598,4	683,9	769,4
20	n_B	65,7	131,3	197,0	262,7	328,3	394,0	459,7	525,3	591,0
	$n_{WT}, \Delta = 0,10$	66,6	133,2	199,8	266,4	333,0	399,7	466,3	532,9	599,5
	$n_{WT}, \Delta = 0,25$	69,1	138,1	207,2	276,2	345,3	414,3	483,4	552,4	621,5
	$n_{WT}, \Delta = 0,40$	75,1	150,2	225,2	300,3	375,4	450,5	525,5	600,6	675,7
	n_{HP}	66,5	133,1	199,6	266,2	332,7	399,3	465,8	532,4	598,9

Çizelge 4.16 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 1,0$ için Örneklem Genişlikleri

		$d = 1,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	15,7	31,4	47,0	62,7	78,4	94,1	109,8	125,4	141,1
	n_P	19,3	38,5	57,8	77,1	96,4	115,6	134,9	154,2	173,4
	n_B	16,1	32,2	48,4	64,5	80,6	96,7	112,8	129,0	145,1
	$n_{WT}, \Delta = 0,10$	16,3	32,6	48,9	65,2	81,5	97,8	114,2	130,5	146,8
	$n_{WT}, \Delta = 0,25$	16,8	33,6	50,4	67,2	84,0	100,9	117,7	134,5	151,3
5	$n_{WT}, \Delta = 0,40$	17,9	35,8	53,7	71,6	89,5	107,4	125,3	143,3	161,2
	n_{HP}	15,9	31,8	47,7	63,7	79,6	95,5	111,4	127,3	143,2
	n_P	20,4	40,8	61,2	81,6	102,0	122,4	142,8	163,2	183,6
	n_B	16,3	32,6	48,9	65,2	81,5	97,8	114,2	130,5	146,8
	$n_{WT}, \Delta = 0,10$	16,5	33,1	49,6	66,1	82,6	99,2	115,7	132,2	148,7
10	$n_{WT}, \Delta = 0,25$	17,1	34,2	51,2	68,3	85,4	102,5	119,5	136,6	153,7
	$n_{WT}, \Delta = 0,40$	18,4	36,8	55,3	73,7	92,1	110,5	129,0	147,4	165,8
	n_{HP}	16,2	32,4	48,6	64,8	81,0	97,2	113,4	129,6	145,8
	n_P	21,0	42,0	62,9	83,9	104,9	125,9	146,9	167,8	188,8
	n_B	16,4	32,8	49,2	65,5	81,9	98,3	114,7	131,1	147,5
15	$n_{WT}, \Delta = 0,10$	16,6	33,2	49,8	66,4	83,0	99,6	116,2	132,8	149,4
	$n_{WT}, \Delta = 0,25$	17,2	34,4	51,6	68,8	86,0	103,2	120,4	137,6	154,8
	$n_{WT}, \Delta = 0,40$	18,6	37,3	55,9	74,6	93,2	111,9	130,5	149,1	167,8
	n_{HP}	16,4	32,9	49,3	65,7	82,2	98,6	115,0	131,5	147,9
	n_P	21,4	42,7	64,1	85,5	106,9	128,2	149,6	171,0	192,3
20	n_B	16,4	32,8	49,3	65,7	82,1	98,5	114,9	131,3	147,8
	$n_{WT}, \Delta = 0,10$	16,7	33,3	50,0	66,6	83,3	99,9	116,6	133,2	149,9
	$n_{WT}, \Delta = 0,25$	17,3	34,5	51,8	69,1	86,3	103,6	120,8	138,1	155,4
	$n_{WT}, \Delta = 0,40$	18,8	37,5	56,3	75,1	93,8	112,6	131,4	150,2	168,9
	n_{HP}	16,6	33,3	49,9	66,5	83,2	99,8	116,5	133,1	149,7

Çizelge 4.17 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 1,5$ için Örneklem Genişlikleri

		$d = 1,5$									
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$	
1	n_f	7,0	13,9	20,9	27,9	34,8	41,8	48,8	55,8	62,7	
	n_P	8,6	17,1	25,7	34,3	42,8	51,4	60,0	68,5	77,1	
	n_B	7,2	14,3	21,5	28,7	35,8	43,0	50,1	57,3	64,5	
	5	$n_{WT}, \Delta = 0,10$	7,2	14,5	21,7	29,0	36,2	43,5	50,7	58,0	65,2
		$n_{WT}, \Delta = 0,25$	7,5	14,9	22,4	29,9	37,4	44,8	52,3	59,8	67,2
		$n_{WT}, \Delta = 0,40$	8,0	15,9	23,9	31,8	39,8	47,8	55,7	63,7	71,6
	n_{HP}	7,1	14,1	21,2	28,3	35,4	42,4	49,5	56,6	63,7	
10	n_P	9,1	18,1	27,2	36,3	45,3	54,4	63,5	72,5	81,6	
	n_B	7,2	14,5	21,7	29,0	36,2	43,5	50,7	58,0	65,2	
	10	$n_{WT}, \Delta = 0,10$	7,3	14,7	22,0	29,4	36,7	44,1	51,4	58,8	66,1
		$n_{WT}, \Delta = 0,25$	7,6	15,2	22,8	30,4	37,9	45,5	53,1	60,7	68,3
		$n_{WT}, \Delta = 0,40$	8,2	16,4	24,6	32,8	40,9	49,1	57,3	65,5	73,7
		n_{HP}	7,2	14,4	21,6	28,8	36,0	43,2	50,4	57,6	64,8
15	n_P	9,3	18,6	28,0	37,3	46,6	55,9	65,3	74,6	83,9	
	n_B	7,3	14,6	21,8	29,1	36,4	43,7	51,0	58,3	65,5	
	15	$n_{WT}, \Delta = 0,10$	7,4	14,8	22,1	29,5	36,9	44,3	51,7	59,0	66,4
		$n_{WT}, \Delta = 0,25$	7,6	15,3	22,9	30,6	38,2	45,9	53,5	61,2	68,8
		$n_{WT}, \Delta = 0,40$	8,3	16,6	24,9	33,1	41,4	49,7	58,0	66,3	74,6
		n_{HP}	7,3	14,6	21,9	29,2	36,5	43,8	51,1	58,4	65,7
20	n_P	9,5	19,0	28,5	38,0	47,5	57,0	66,5	76,0	85,5	
	n_B	7,3	14,6	21,9	29,2	36,5	43,8	51,1	58,4	65,7	
	20	$n_{WT}, \Delta = 0,10$	7,4	14,8	22,2	29,6	37,0	44,4	51,8	59,2	66,6
		$n_{WT}, \Delta = 0,25$	7,7	15,3	23,0	30,7	38,4	46,0	53,7	61,4	69,1
		$n_{WT}, \Delta = 0,40$	8,3	16,7	25,0	33,4	41,7	50,1	58,4	66,7	75,1
		n_{HP}	7,4	14,8	22,2	29,6	37,0	44,4	51,8	59,2	66,5

Çizelge 4.18 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 2,0$ için Örneklem Genişlikleri

		$d = 2,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	3,9	7,8	11,8	15,7	19,6	23,5	27,4	31,4	35,3
	n_P	4,8	9,6	14,5	19,3	24,1	28,9	33,7	38,5	43,4
	n_B	4,0	8,1	12,1	16,1	20,1	24,2	28,2	32,2	36,3
	$n_{WT}, \Delta = 0,10$	4,1	8,2	12,2	16,3	20,4	24,5	28,5	32,6	36,7
	$n_{WT}, \Delta = 0,25$	4,2	8,4	12,6	16,8	21,0	25,2	29,4	33,6	37,8
5	$n_{WT}, \Delta = 0,40$	4,5	9,0	13,4	17,9	22,4	26,9	31,3	35,8	40,3
	n_{HP}	4,0	8,0	11,9	15,9	19,9	23,9	27,9	31,8	35,8
	n_P	5,1	10,2	15,3	20,4	25,5	30,6	35,7	40,8	45,9
	n_B	4,1	8,2	12,2	16,3	20,4	24,5	28,5	32,6	36,7
	$n_{WT}, \Delta = 0,10$	4,1	8,3	12,4	16,5	20,7	24,8	28,9	33,1	37,2
10	$n_{WT}, \Delta = 0,25$	4,3	8,5	12,8	17,1	21,3	25,6	29,9	34,2	38,4
	$n_{WT}, \Delta = 0,40$	4,6	9,2	13,8	18,4	23,0	27,6	32,2	36,8	41,5
	n_{HP}	4,0	8,1	12,1	16,2	20,2	24,3	28,3	32,4	36,4
	n_P	5,2	10,5	15,7	21,0	26,2	31,5	36,7	42,0	47,2
	n_B	4,1	8,2	12,3	16,4	20,5	24,6	28,7	32,8	36,9
15	$n_{WT}, \Delta = 0,10$	4,2	8,3	12,5	16,6	20,8	24,9	29,1	33,2	37,4
	$n_{WT}, \Delta = 0,25$	4,3	8,6	12,9	17,2	21,5	25,8	30,1	34,4	38,7
	$n_{WT}, \Delta = 0,40$	4,7	9,3	14,0	18,6	23,3	28,0	32,6	37,3	41,9
	n_{HP}	4,1	8,2	12,3	16,4	20,5	24,6	28,8	32,9	37,0
	n_P	5,3	10,7	16,0	21,4	26,7	32,1	37,4	42,7	48,1
20	n_B	4,1	8,2	12,3	16,4	20,5	24,6	28,7	32,8	36,9
	$n_{WT}, \Delta = 0,10$	4,2	8,3	12,5	16,7	20,8	25,0	29,1	33,3	37,5
	$n_{WT}, \Delta = 0,25$	4,3	8,6	12,9	17,3	21,6	25,9	30,2	34,5	38,8
	$n_{WT}, \Delta = 0,40$	4,7	9,4	14,1	18,8	23,5	28,2	32,8	37,5	42,2
	n_{HP}	4,2	8,3	12,5	16,6	20,8	25,0	29,1	33,3	37,4

Çizelge 4.19 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 2,5$ için Örneklem Genişlikleri

		$d = 2,5$									
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$	
1	n_f	2,5	5,0	7,5	10,0	12,5	15,1	17,6	20,1	22,6	
	n_P	3,1	6,2	9,2	12,3	15,4	18,5	21,6	24,7	27,7	
	n_B	2,6	5,2	7,7	10,3	12,9	15,5	18,1	20,6	23,2	
	5	$n_{WT}, \Delta = 0,10$	2,6	5,2	7,8	10,4	13,0	15,7	18,3	20,9	23,5
		$n_{WT}, \Delta = 0,25$	2,7	5,4	8,1	10,8	13,4	16,1	18,8	21,5	24,2
		$n_{WT}, \Delta = 0,40$	2,9	5,7	8,6	11,5	14,3	17,2	20,1	22,9	25,8
	n_{HP}	2,5	5,1	7,6	10,2	12,7	15,3	17,8	20,4	22,9	
10	n_P	3,3	6,5	9,8	13,1	16,3	19,6	22,8	26,1	29,4	
	n_B	2,6	5,2	7,8	10,4	13,0	15,7	18,3	20,9	23,5	
	10	$n_{WT}, \Delta = 0,10$	2,6	5,3	7,9	10,6	13,2	15,9	18,5	21,2	23,8
		$n_{WT}, \Delta = 0,25$	2,7	5,5	8,2	10,9	13,7	16,4	19,1	21,9	24,6
		$n_{WT}, \Delta = 0,40$	2,9	5,9	8,8	11,8	14,7	17,7	20,6	23,6	26,5
		n_{HP}	2,6	5,2	7,8	10,4	13,0	15,5	18,1	20,7	23,3
15	n_P	3,4	6,7	10,1	13,4	16,8	20,1	23,5	26,9	30,2	
	n_B	2,6	5,2	7,9	10,5	13,1	15,7	18,4	21,0	23,6	
	15	$n_{WT}, \Delta = 0,10$	2,7	5,3	8,0	10,6	13,3	15,9	18,6	21,3	23,9
		$n_{WT}, \Delta = 0,25$	2,8	5,5	8,3	11,0	13,8	16,5	19,3	22,0	24,8
		$n_{WT}, \Delta = 0,40$	3,0	6,0	8,9	11,9	14,9	17,9	20,9	23,9	26,8
		n_{HP}	2,6	5,3	7,9	10,5	13,1	15,8	18,4	21,0	23,7
20	n_P	3,4	6,8	10,3	13,7	17,1	20,5	23,9	27,4	30,8	
	n_B	2,6	5,3	7,9	10,5	13,1	15,8	18,4	21,0	23,6	
	20	$n_{WT}, \Delta = 0,10$	2,7	5,3	8,0	10,7	13,3	16,0	18,7	21,3	24,0
		$n_{WT}, \Delta = 0,25$	2,8	5,5	8,3	11,0	13,8	16,6	19,3	22,1	24,9
		$n_{WT}, \Delta = 0,40$	3,0	6,0	9,0	12,0	15,0	18,0	21,0	24,0	27,0
		n_{HP}	2,7	5,3	8,0	10,6	13,3	16,0	18,6	21,3	24,0

Çizelge 4.20 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 3,0$ için Örneklem Genişlikleri

		$d = 3,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	1,7	3,5	5,2	7,0	8,7	10,5	12,2	13,9	15,7
	n_P	2,1	4,3	6,4	8,6	10,7	12,8	15,0	17,1	19,3
	n_B	1,8	3,6	5,4	7,2	9,0	10,7	12,5	14,3	16,1
	$n_{WT}, \Delta = 0,10$	1,8	3,6	5,4	7,2	9,1	10,9	12,7	14,5	16,3
	$n_{WT}, \Delta = 0,25$	1,9	3,7	5,6	7,5	9,3	11,2	13,1	14,9	16,8
5	$n_{WT}, \Delta = 0,40$	2,0	4,0	6,0	8,0	9,9	11,9	13,9	15,9	17,9
	n_{HP}	1,8	3,5	5,3	7,1	8,8	10,6	12,4	14,1	15,9
	n_P	2,3	4,5	6,8	9,1	11,3	13,6	15,9	18,1	20,4
	n_B	1,8	3,6	5,4	7,2	9,1	10,9	12,7	14,5	16,3
	$n_{WT}, \Delta = 0,10$	1,8	3,7	5,5	7,3	9,2	11,0	12,9	14,7	16,5
10	$n_{WT}, \Delta = 0,25$	1,9	3,8	5,7	7,6	9,5	11,4	13,3	15,2	17,1
	$n_{WT}, \Delta = 0,40$	2,0	4,1	6,1	8,2	10,2	12,3	14,3	16,4	18,4
	n_{HP}	1,8	3,6	5,4	7,2	9,0	10,8	12,6	14,4	16,2
	n_P	2,3	4,7	7,0	9,3	11,7	14,0	16,3	18,6	21,0
	n_B	1,8	3,6	5,5	7,3	9,1	10,9	12,7	14,6	16,4
15	$n_{WT}, \Delta = 0,10$	1,8	3,7	5,5	7,4	9,2	11,1	12,9	14,8	16,6
	$n_{WT}, \Delta = 0,25$	1,9	3,8	5,7	7,6	9,6	11,5	13,4	15,3	17,2
	$n_{WT}, \Delta = 0,40$	2,1	4,1	6,2	8,3	10,4	12,4	14,5	16,6	18,6
	n_{HP}	1,8	3,7	5,5	7,3	9,1	11,0	12,8	14,6	16,4
	n_P	2,4	4,7	7,1	9,5	11,9	14,2	16,6	19,0	21,4
20	n_B	1,8	3,6	5,5	7,3	9,1	10,9	12,8	14,6	16,4
	$n_{WT}, \Delta = 0,10$	1,9	3,7	5,6	7,4	9,3	11,1	13,0	14,8	16,7
	$n_{WT}, \Delta = 0,25$	1,9	3,8	5,8	7,7	9,6	11,5	13,4	15,3	17,3
	$n_{WT}, \Delta = 0,40$	2,1	4,2	6,3	8,3	10,4	12,5	14,6	16,7	18,8
	n_{HP}	1,8	3,7	5,5	7,4	9,2	11,1	12,9	14,8	16,6

Çizelge 4.21 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 0,5$ için Örneklem Genişlikleri

		$d = 0,5$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	118,9	237,8	356,7	475,6	594,4	713,3	832,2	951,1	1070,0
5	n_P	139,1	278,2	417,3	556,4	695,5	834,6	973,7	1112,8	1251,9
	n_B	120,6	241,1	361,7	482,2	602,8	723,3	843,9	964,4	1085,0
10	n_P	145,3	290,6	435,8	581,1	726,4	871,7	1017,0	1162,3	1307,5
	n_B	121,5	243,0	364,5	486,0	607,5	729,0	850,5	972,0	1093,5
15	n_P	148,4	296,7	445,1	593,5	741,9	890,2	1038,6	1187,0	1335,4
	n_B	122,0	244,0	365,9	487,9	609,9	731,9	853,9	975,8	1097,8
20	n_P	150,3	300,5	450,8	601,1	751,4	901,6	1051,9	1202,2	1352,5
	n_B	122,3	244,7	367,0	489,3	611,7	734,0	856,4	978,7	1101,0

Çizelge 4.22 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 1,0$ için Örneklem Genişlikleri

		$d = 1,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	29,7	59,4	89,2	118,9	148,6	178,3	208,1	237,8	267,5
5	n_P	34,8	69,5	104,3	139,1	173,9	208,6	243,4	278,2	313,0
	n_B	30,1	60,3	90,4	120,6	150,7	180,8	211,0	241,1	271,2
10	n_P	36,3	72,6	109,0	145,3	181,6	217,9	254,2	290,6	326,9
	n_B	30,4	60,8	91,1	121,5	151,9	182,3	212,6	243,0	273,4
15	n_P	37,1	74,2	111,3	148,4	185,5	222,6	259,7	296,7	333,8
	n_B	30,5	61,0	91,5	122,0	152,5	183,0	213,5	244,0	274,5
20	n_P	37,6	75,1	112,7	150,3	187,8	225,4	263,0	300,5	338,1
	n_B	30,6	61,2	91,8	122,3	152,9	183,5	214,1	244,7	275,3

Çizelge 4.23 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 1,5$ için Örneklem Genişlikleri

		$d = 1,5$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	13,2	26,4	39,6	52,8	66,0	79,3	92,5	105,7	118,9
5	n_P	15,5	30,9	46,4	61,8	77,3	92,7	108,2	123,6	139,1
	n_B	13,4	26,8	40,2	53,6	67,0	80,4	93,8	107,2	120,6
10	n_P	16,1	32,3	48,4	64,6	80,7	96,9	113,0	129,1	145,3
	n_B	13,5	27,0	40,5	54,0	67,5	81,0	94,5	108,0	121,5
15	n_P	16,5	33,0	49,5	65,9	82,4	98,9	115,4	131,9	148,4
	n_B	13,6	27,1	40,7	54,2	67,8	81,3	94,9	108,4	122,0
20	n_P	16,7	33,4	50,1	66,8	83,5	100,2	116,9	133,6	150,3
	n_B	13,6	27,2	40,8	54,4	68,0	81,6	95,2	108,7	122,3

Çizelge 4.24 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 2,0$ için Örneklem Genişlikleri

		$d = 2,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	7,4	14,9	22,3	29,7	37,2	44,6	52,0	59,4	66,9
5	n_P	8,7	17,4	26,1	34,8	43,5	52,2	60,9	69,5	78,2
	n_B	7,5	15,1	22,6	30,1	37,7	45,2	52,7	60,3	67,8
10	n_P	9,1	18,2	27,2	36,3	45,4	54,5	63,6	72,6	81,7
	n_B	7,6	15,2	22,8	30,4	38,0	45,6	53,2	60,8	68,3
15	n_P	9,3	18,5	27,8	37,1	46,4	55,6	64,9	74,2	83,5
	n_B	7,6	15,2	22,9	30,5	38,1	45,7	53,4	61,0	68,6
20	n_P	9,4	18,8	28,2	37,6	47,0	56,4	65,7	75,1	84,5
	n_B	7,6	15,3	22,9	30,6	38,2	45,9	53,5	61,2	68,8

Çizelge 4.25 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 2,5$ için Örneklem Genişlikleri

		$d = 2,5$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	4,8	9,5	14,3	19,0	23,8	28,5	33,3	38,0	42,8
5	n_P	5,6	11,1	16,7	22,3	27,8	33,4	38,9	44,5	50,1
	n_B	4,8	9,6	14,5	19,3	24,1	28,9	33,8	38,6	43,4
10	n_P	5,8	11,6	17,4	23,2	29,1	34,9	40,7	46,5	52,3
	n_B	4,9	9,7	14,6	19,4	24,3	29,2	34,0	38,9	43,7
15	n_P	5,9	11,9	17,8	23,7	29,7	35,6	41,5	47,5	53,4
	n_B	4,9	9,8	14,6	19,5	24,4	29,3	34,2	39,0	43,9
20	n_P	6,0	12,0	18,0	24,0	30,1	36,1	42,1	48,1	54,1
	n_B	4,9	9,8	14,7	19,6	24,5	29,4	34,3	39,1	44,0

Çizelge 4.26 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 3,0$ için Örneklem Genişlikleri

		$d = 3,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	3,3	6,6	9,9	13,2	16,5	19,8	23,1	26,4	29,7
5	n_P	3,9	7,7	11,6	15,5	19,3	23,2	27,0	30,9	34,8
	n_B	3,3	6,7	10,0	13,4	16,7	20,1	23,4	26,8	30,1
10	n_P	4,0	8,1	12,1	16,1	20,2	24,2	28,2	32,3	36,3
	n_B	3,4	6,8	10,1	13,5	16,9	20,3	23,6	27,0	30,4
15	n_P	4,1	8,2	12,4	16,5	20,6	24,7	28,9	33,0	37,1
	n_B	3,4	6,8	10,2	13,6	16,9	20,3	23,7	27,1	30,5
20	n_P	4,2	8,3	12,5	16,7	20,9	25,0	29,2	33,4	37,6
	n_B	3,4	6,8	10,2	13,6	17,0	20,4	23,8	27,2	30,6

Çizelge 4.27 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 0,5$ için Örneklem Genişlikleri

		$d = 0,5$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	93,3	186,6	279,9	373,2	466,5	559,8	653,1	746,4	839,7
5	n_P	110,7	221,5	332,2	443,0	553,7	664,5	775,2	886,0	996,7
	n_B	94,7	189,4	284,1	378,8	473,5	568,2	662,9	757,6	852,3
10	n_P	116,0	231,9	347,9	463,9	579,8	695,8	811,8	927,8	1043,7
	n_B	95,5	191,1	286,6	382,1	477,7	573,2	668,8	764,3	859,8
15	n_P	118,7	237,3	356,0	474,7	593,4	712,0	830,7	949,4	1068,1
	n_B	95,9	191,8	287,7	383,6	479,6	575,5	671,4	767,3	863,2
20	n_P	120,4	240,9	361,3	481,8	602,2	722,7	843,1	963,6	1084,0
	n_B	96,1	192,2	288,3	384,4	480,5	576,6	672,7	768,8	864,9

Çizelge 4.28 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 1,0$ için Örneklem Genişlikleri

		$d = 1,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	23,3	46,6	70,0	93,3	116,6	139,9	163,3	186,6	209,9
5	n_P	27,7	55,4	83,1	110,7	138,4	166,1	193,8	221,5	249,2
	n_B	23,7	47,3	71,0	94,7	118,4	142,0	165,7	189,4	213,1
10	n_P	29,0	58,0	87,0	116,0	145,0	174,0	202,9	231,9	260,9
	n_B	23,9	47,8	71,7	95,5	119,4	143,3	167,2	191,1	215,0
15	n_P	29,7	59,3	89,0	118,7	148,3	178,0	207,7	237,3	267,0
	n_B	24,0	48,0	71,9	95,9	119,9	143,9	167,8	191,8	215,8
20	n_P	30,1	60,2	90,3	120,4	150,6	180,7	210,8	240,9	271,0
	n_B	24,0	48,0	72,1	96,1	120,1	144,1	168,2	192,2	216,2

Çizelge 4.29 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 1,5$ için Örneklem Genişlikleri

		$d = 1,5$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	10,4	20,7	31,1	41,5	51,8	62,2	72,6	82,9	93,3
5	n_P	12,3	24,6	36,9	49,2	61,5	73,8	86,1	98,4	110,7
	n_B	10,5	21,0	31,6	42,1	52,6	63,1	73,7	84,2	94,7
10	n_P	12,9	25,8	38,7	51,5	64,4	77,3	90,2	103,1	116,0
	n_B	10,6	21,2	31,8	42,5	53,1	63,7	74,3	84,9	95,5
15	n_P	13,2	26,4	39,6	52,7	65,9	79,1	92,3	105,5	118,7
	n_B	10,7	21,3	32,0	42,6	53,3	63,9	74,6	85,3	95,9
20	n_P	13,4	26,8	40,1	53,5	66,9	80,3	93,7	107,1	120,4
	n_B	10,7	21,4	32,0	42,7	53,4	64,1	74,7	85,4	96,1

Çizelge 4.30 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 2,0$ için Örneklem Genişlikleri

		$d = 2,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	5,8	11,7	17,5	23,3	29,2	35,0	40,8	46,6	52,5
5	n_P	6,9	13,8	20,8	27,7	34,6	41,5	48,5	55,4	62,3
	n_B	5,9	11,8	17,8	23,7	29,6	35,5	41,4	47,3	53,3
10	n_P	7,2	14,5	21,7	29,0	36,2	43,5	50,7	58,0	65,2
	n_B	6,0	11,9	17,9	23,9	29,9	35,8	41,8	47,8	53,7
15	n_P	7,4	14,8	22,3	29,7	37,1	44,5	51,9	59,3	66,8
	n_B	6,0	12,0	18,0	24,0	30,0	36,0	42,0	48,0	53,9
20	n_P	7,5	15,1	22,6	30,1	37,6	45,2	52,7	60,2	67,8
	n_B	6,0	12,0	18,0	24,0	30,0	36,0	42,0	48,0	54,1

Çizelge 4.31 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 2,5$ için Örneklem Genişlikleri

		$d = 2,5$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	3,7	7,5	11,2	14,9	18,7	22,4	26,1	29,9	33,6
5	n_P	4,4	8,9	13,3	17,7	22,1	26,6	31,0	35,4	39,9
	n_B	3,8	7,6	11,4	15,2	18,9	22,7	26,5	30,3	34,1
10	n_P	4,6	9,3	13,9	18,6	23,2	27,8	32,5	37,1	41,7
	n_B	3,8	7,6	11,5	15,3	19,1	22,9	26,8	30,6	34,4
15	n_P	4,7	9,5	14,2	19,0	23,7	28,5	33,2	38,0	42,7
	n_B	3,8	7,7	11,5	15,3	19,2	23,0	26,9	30,7	34,5
20	n_P	4,8	9,6	14,5	19,3	24,1	28,9	33,7	38,5	43,4
	n_B	3,8	7,7	11,5	15,4	19,2	23,1	26,9	30,8	34,6

Çizelge 4.32 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 3,0$ için Örneklem Genişlikleri

		$d = 3,0$								
K	n	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 3$	$\sigma^2 = 4$	$\sigma^2 = 5$	$\sigma^2 = 6$	$\sigma^2 = 7$	$\sigma^2 = 8$	$\sigma^2 = 9$
1	n_f	2,6	5,2	7,8	10,4	13,0	15,5	18,1	20,7	23,3
5	n_P	3,1	6,2	9,2	12,3	15,4	18,5	21,5	24,6	27,7
	n_B	2,6	5,3	7,9	10,5	13,2	15,8	18,4	21,0	23,7
10	n_P	3,2	6,4	9,7	12,9	16,1	19,3	22,5	25,8	29,0
	n_B	2,7	5,3	8,0	10,6	13,3	15,9	18,6	21,2	23,9
15	n_P	3,3	6,6	9,9	13,2	16,5	19,8	23,1	26,4	29,7
	n_B	2,7	5,3	8,0	10,7	13,3	16,0	18,6	21,3	24,0
20	n_P	3,3	6,7	10,0	13,4	16,7	20,1	23,4	26,8	30,1
	n_B	2,7	5,3	8,0	10,7	13,3	16,0	18,7	21,4	24,0

5. TARTIŞMA

Sabit örneklem genişliği testlerinde dikkate alınan kritik değer (c_K) için $c_K = C_P(K, \alpha)$, $c_K = C_B(K, \alpha)$, $c_K = C_{WT}(K, \alpha, \Delta)$ ve $c_K = C_{HP}(K, \alpha)$ olduğundan, sırasıyla $C_P(K, \alpha)$, $C_B(K, \alpha)$, $C_{WT}(K, \alpha, \Delta)$ ve $C_{HP}(K, \alpha)$ değerlerini $C_P(1, \alpha)$, $C_B(1, \alpha)$, $C_{WT}(1, \alpha, \Delta)$ ve $C_{HP}(1, \alpha)$ değerleriyle karşılaştırmak, grup ardışık test tiplerinde son analizde hesaplanan Z_K test istatistiğini karşılaştırmak için dikkate alınan kritik değerlerin, sabit örneklem genişliği testinde dikkate alınan kritik değerden ne kadar büyük olduğunu gösterecektir (3).

$\alpha = 0,05$ I. tip hata düzeyinde, sabit örneklem genişliği testinde H_0 hipotezini reddetmek için dikkate alınan kritik değer $c_K = 1,96$ 'dır. Çizelge 4.1, Çizelge 4.2, Çizelge 4.3 ve Çizelge 4.4 incelendiğinde, $\alpha = 0,05$ I. tip hata düzeyinde, sabit örneklem genişliği testinde dikkate alınan kritik değere en yakın kritik değerlerin Haybittle-Peto testi son analizinde dikkate alındığı, Pocock testi son analizinde dikkate alınan kritik değerlerin ise en yüksek farka sahip olduğu görülmektedir. Bu durum grup ardışık testlerde yapılacak toplam analiz sayısına göre de değişiklik göstermemektedir. $\alpha = 0,05$ I. tip hata düzeyinde, $K = 5$, $K = 10$, $K = 15$ ve $K = 20$ toplam analiz sayısı ile uygulanan grup ardışık test tiplerinin her bir ara analiz ve son analizde dikkate aldıkları tüm kritik değerler sırasıyla Çizelge 4.1, Çizelge 4.2, Çizelge 4.3 ve Çizelge 4.4'te verilmiş, ayrıca Şekil 5.1, Şekil 5.2, Şekil 5.3 ve Şekil 5.4'te de grafik üzerinde gösterilmiştir.

Eğer grup ardışık deneme düzeninin uygulandığı bir deneme boyunca yapılan ara analizlerde H_0 hipotezini reddedecek bir sonuca ulaşılamaz ve deneme son analize kadar devam ederse, sabit örneklem genişliği deneme düzeni kullanılan bir denemeye göre en az farkla karar verecek olan grup ardışık test tipi Haybittle-Peto testidir. Daha sonra O'Brien & Fleming testinin en yakın kritik değere sahip olduğu görülürken, bunu

sırasıyla $\Delta = 0,10$, $\Delta = 0,25$ ve $\Delta = 0,40$ parametrelili Wang & Tsiatis testleri izlemektedir. Son analizde dikkate alınan kritik değerler bakımından, sabit örneklem genişliği testinde kullanılan kritik değere en uzak kritik değer ise Pocock testinde gözlenmektedir. Wang & Tsiatis testinde $\Delta = 0,00$ olarak dikkate alındığında O'Brien & Fleming testiyle ve $\Delta = 0,50$ olarak dikkate alındığında ise Pocock testiyle aynı olduğundan bu parametre değerleri için gözlenen değerler de bu sıralamayı etkilememektedir. Bu durum grup ardışık test tiplerinin son analizde H_0 hipotezini reddetmek için dikkate aldıkları kritik değerler bakımından birbirlerine göre durumlarını da göstermektedir ve sıralama yine aynı şekildedir. Ancak bu sıralama, eğer grup ardışık deneme düzeninin uygulandığı bir deneme boyunca yapılan ara analizlerde H_0 hipotezini reddedecek bir sonuca ulaşamaz ve deneme son analize kadar devam ederse ortaya çıkmaktadır.

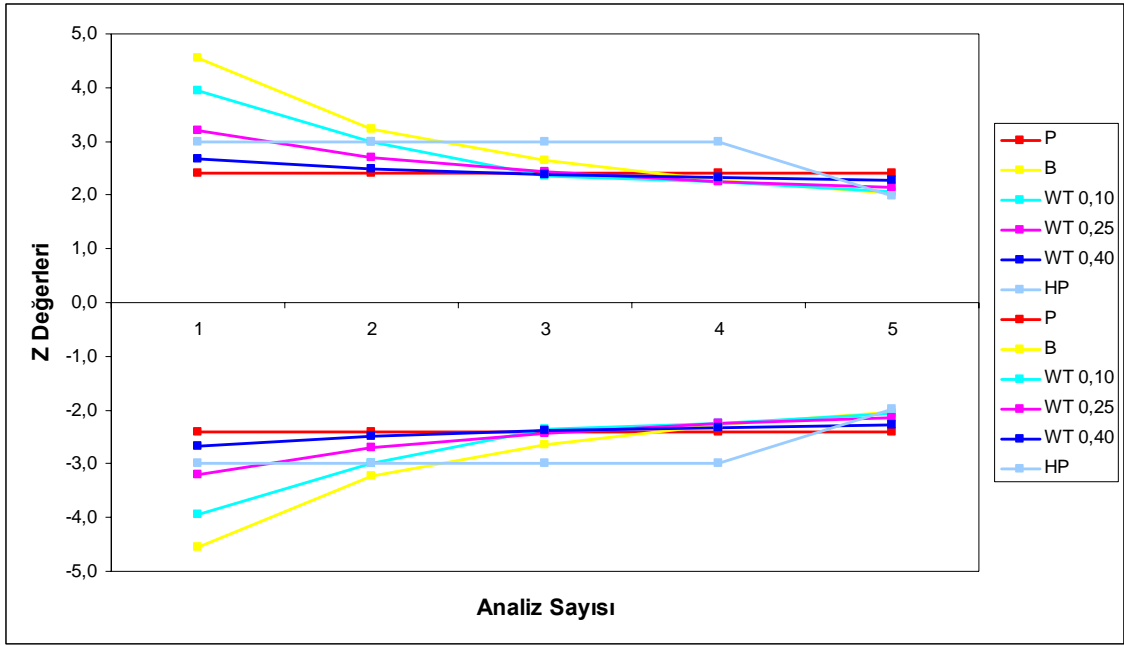
Grup ardışık test tiplerinin ara analizlerde dikkate aldıkları kritik değerler incelendiğinde ise bu durum biraz farklılık göstermekte ve sıralama her bir ara analizde farklı olarak ortaya çıkmaktadır. Yine Çizelge 4.1, Çizelge 4.2, Çizelge 4.3 ve Çizelge 4.4'te verilen, sırasıyla $K = 5$, $K = 10$, $K = 15$ ve $K = 20$ toplam analiz sayısı ve $\alpha = 0,05$ I. tip hata düzeyinde uygulanan grup ardışık test tiplerinin her bir ara analiz ve son analizde dikkate aldıkları tüm kritik değerler incelendiğinde test tiplerinin birbirine göre durumları da görülebilmektedir. Şekil 5.1, Şekil 5.2, Şekil 5.3 ve Şekil 5.4 ise bu farklılıkların grafiksel gösterimini sağlamaktadır. Şekiller incelendiğinde, ilk analizler için O'Brien & Fleming testinde dikkate alınan kritik değerler diğer test tiplerine göre daha yüksek olduğundan, bu test tipinde H_0 hipotezini erken aşamalarda reddetmek oldukça zor görünmektedir. Ayrıca yapılması planlanan toplam analiz sayısı (K) arttıkça ilk analizler için $\sqrt{(K/k)}$ değeri artacağından $C_B(K, \alpha) \cdot \sqrt{(K/k)}$ değerleri de artmakta ve diğer testlere göre olan farklılık K arttıkça daha da yüksek olmaktadır. Buna karşın son (final) analize yaklaştıkça $\sqrt{(K/k)}$ değeri küçüleceğinden $C_B(K, \alpha) \cdot \sqrt{(K/k)}$ değerlerindeki artış miktarı da azalmaya başlamakta ve sona yakın ara analizlerde diğer testlerden daha küçük kritik değerler ortaya çıkmaktadır. Benzer şekilde Wang & Tsiatis testinde de bu durum gözlenmekte, yapılması planlanan toplam

analiz sayısı (K) arttıkça ilk analizler için $(k/K)^{\Delta-1/2}$ değeri artacağından $C_{WT}(K, \alpha, \Delta) \cdot (k/K)^{\Delta-1/2}$ değerleri de artmakta ve K arttıkça diğer testlere göre daha yüksek farklılıklar ortaya çıkmaktadır. Bunun yanında yüksek Δ seçimleri için bu artış az miktardaiken Δ değerleri azaldıkça Wang & Tsiatis testi O'Brien & Fleming testine yaklaştığından yine ilk analizlerde H_0 hipotezini reddetmek için diğer testlere göre daha yüksek kritik değerlere sahip olmakta ve sona yakın ara analizlerde de yine diğer testlerden daha küçük kritik değerler dikkate almaktadır.

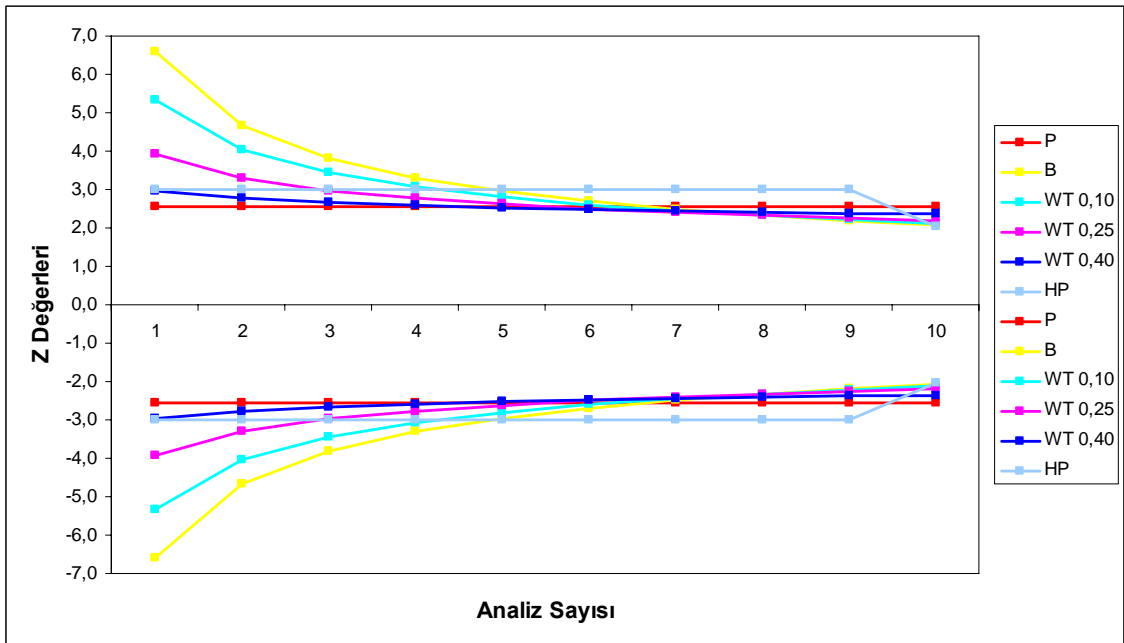
Örneğin; $\alpha = 0,05$ I. tip hata düzeyinde, toplam analiz sayısı $K = 5$ olduğunda, Pocock testi için tüm ara analizlerde dikkate alınacak kritik değerler 2,413'te sabittir. O'Brien & Fleming testinin ara analizlerde dikkate alacağı kritik değerler ise $c_k = C_B(5, 0,05) \cdot \sqrt{(5/k)} = 2,040 \cdot \sqrt{(5/k)}$ formülünden sırasıyla 4,562, 3,226, 2,634, 2,281 ve 2,040 olarak hesaplanmaktadır. $\Delta = 0,25$ olan Wang & Tsiatis testi için kritik değerler ise yine $c_k = C_{WT}(5, 0,05, 0,25) \cdot (k/5)^{0,25-1/2} = 2,136 \cdot (k/5)^{-0,25}$ formülünden sırasıyla 3,194, 2,686, 2,427, 2,259 ve 2,136 şeklinde bulunmaktadır. Haybittle-Peto testi, son (final) analize kadar yüksek kritik değerler ($c_k = 3$) kullanmasına rağmen, bu değerler ilk bir kaç O'Brien & Fleming ara analizinde kullanılan kritik değerlerden (örneğin burada $c_1 = 4,562$ ve $c_2 = 3,226$) daha küçüktür.

Başka bir örnek olarak; $\alpha = 0,05$ I. tip hata olasılığı ve $K = 10$ olan testlerde kritik değerler incelendiğinde, Haybittle-Peto testi ve O'Brien & Fleming testi (aynı zamanda $\Delta = 0,00$ olan Wang & Tsiatis testidir), yine ilk ara analizlerde çok geniş sınırlara sahiptir. Son (final) analizde dikkate alınan kritik değer açısından ise sabit örneklem genişliği testinde kullanılan 1,96'ya en yakın değerlere ise yine bu testler sahiptir. Pocock testi (aynı zamanda $\Delta = 0,50$ olan Wang & Tsiatis testidir) ise $c_k = 2,555$ sabit kritik değerine sahiptir. $\Delta = 0,25$ olan Wang & Tsiatis testinin kritik değerleri ise bu değerlerin arasında kalmakta ve son analizdeki kritik değer olan $c_{10} = 2,199$ 'a doğru düzgün şekilde azalmaktadır.

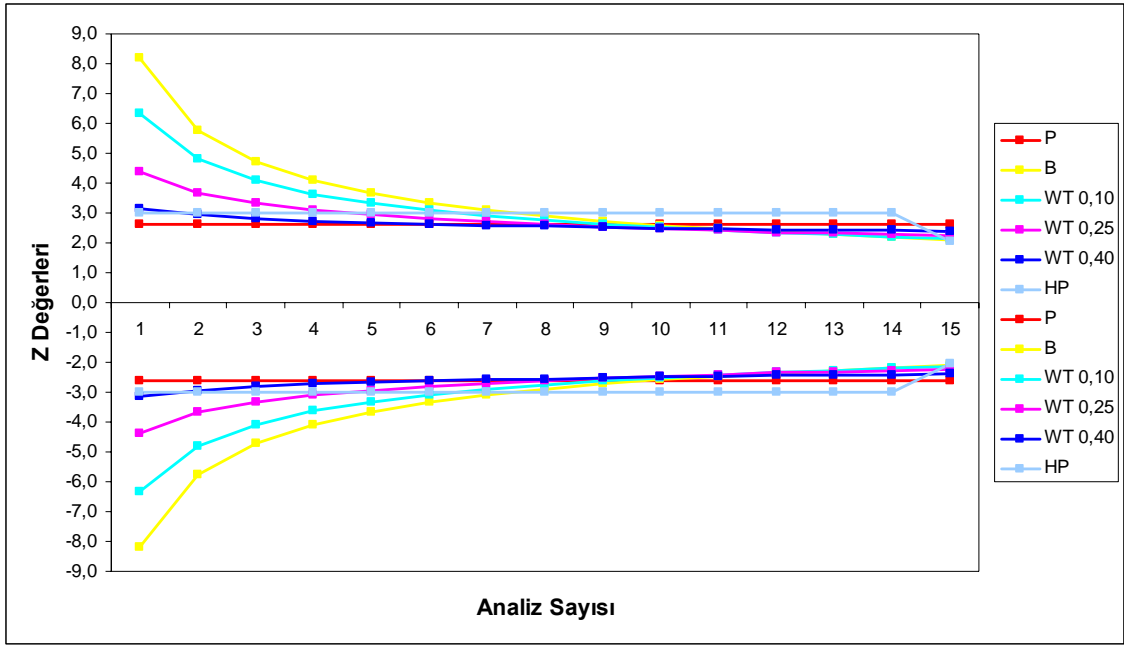
Yapılan tüm kombinasyonlarda elde edilen sonuçlar genel olarak incelendiğinde verilen örneklere benzer sonuçlar ortaya çıkmaktadır. O'Brien & Fleming testi Pocock testi ile karşılaştırıldığında, $K = 5$ olduğunda ilk 3 ara analizde Pocock testinin daha küçük etki büyüklüğünü tespit edebilecek kritik değerlere sahip olduğu, 4. ara analiz ve son analizde ise O'Brien & Fleming testinin daha küçük etki büyüklüklerini tespit edebilecek duruma geldiği görülmektedir. $K = 10$ olduğunda bu değişim 7. ara analizden itibaren, $K = 15$ olduğunda 10. ara analizden itibaren ve $K = 20$ olduğunda ise 13. ara analizden itibaren ortaya çıkmaktadır. Wang & Tsatis testi 0,00-0,50 arasındaki farklı Δ değeri seçimlerinde, O'Brien & Fleming testi ile Pocock testi arasında kalan değerlerle sonuçlandığından bu durum bu test için de çok farklı olmamaktadır ve Δ değeri arttıkça ilk ara analizde dikkate alınan kritik değerler Pocock testine daha çok yaklaşmakta ancak sona yakın analizlere kadar yine Pocock testine göre daha büyük kalmaktadır. Farklı Δ seçimlerine göre değişmekle birlikte, $K = 5$ olduğunda ilk 2-3 ara analizde Pocock testi daha küçük etki büyüklüğünü tespit edebilecek kritik değerlere sahipken, 3.-4. ara analiz ve son analizde ise Wang & Tsatis testinin daha küçük etki büyüklüklerini tespit edebilecek duruma geldiği görülmektedir. $K = 10$ olduğunda bu değişim 5.-7. ara analizlerden itibaren, $K = 15$ olduğunda 7.-10. ara analizlerden itibaren ve $K = 20$ olduğunda ise 8.-11. ara analizlerden itibaren başlamaktadır. Haybittle-Peto testinde toplam analiz sayısı ne olursa olsun tüm ara analizler için dikkate alınan kritik değerler $c_1 = \dots = c_{K-1} = 3$ olduğundan başlangıçta Pocock testi hariç tüm testlerin dikkate aldığı kritik değerlerden daha küçük kritik değerler kullanılırken, daha sonraki ara analizlerde tüm testler Haybittle-Peto testinden daha küçük kritik değerler kullanmaktadır. $K = 5$ olduğunda O'Brien & Fleming testi hariç tüm testler 2. ara analizden itibaren Haybittle-Peto testine göre daha küçük kritik değerler kullanmaya başlamaktadır. $K = 10$ olduğunda bu değişim 2.-5. ara analizlerden itibaren, $K = 15$ olduğunda 3.-8. ara analizlerden itibaren ve $K = 20$ olduğunda ise 3.-11. ara analizlerden itibaren başlamaktadır. Son analize gelindiğinde ise toplam analiz sayısı ne olursa olsun, en küçük kritik değere Haybittle-Peto testi sahip olmaktadır.



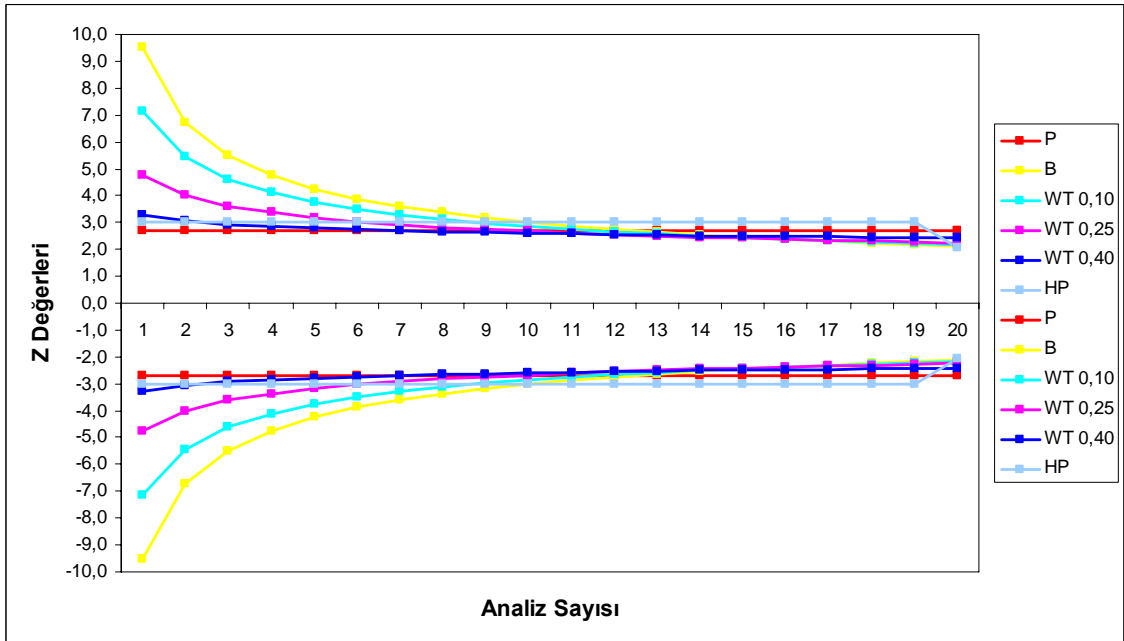
Şekil 5.1 $\alpha = 0,05$ ve $K = 5$ için Kritik Değerler



Şekil 5.2 $\alpha = 0,05$ ve $K = 10$ için Kritik Değerler



Şekil 5.3 $\alpha = 0,05$ ve $K = 15$ için Kritik Değerler



Şekil 5.4 $\alpha = 0,05$ ve $K = 20$ için Kritik Değerler

Bölüm 2.6.3.4'te açıklandığı gibi $\alpha = 0,05$ düzeyinden daha düşük bir I. tip hata olasılığı sağlanamadığından Haybittle-Peto testinde H_0 hipotezini reddetmek için dikkate alınan kritik değerler sadece $\alpha = 0,05$ I. tip hata düzeyinde hesaplanabilmektedir. Wang & Tsiatis testleri ise farklı Δ değerleri için Pocock ve O'Brien & Fleming testlerinin arasında kalan durumları gösterdiğinden $\alpha = 0,01$ I. tip hata düzeyinde sadece Pocock ve O'Brien & Fleming testleri birbirileriyle ve sabit örneklem genişliği deneme düzeni testiyle karşılaştırılmıştır.

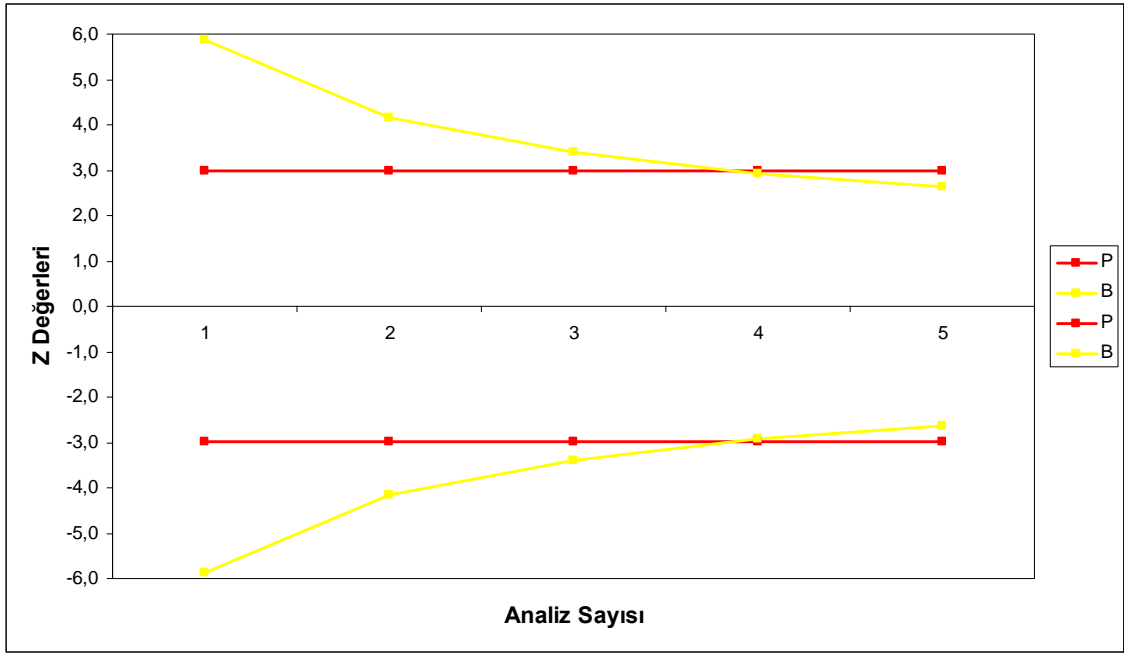
$\alpha = 0,01$ I. tip hata düzeyinde, sabit örneklem genişliği testinde H_0 hipotezini reddetmek için dikkate alınan kritik değer $c_K = 2,575$ 'tir. Çizelge 4.5, Çizelge 4.6, Çizelge 4.7 ve Çizelge 4.8 incelendiğinde, $\alpha = 0,01$ I. tip hata düzeyinde, sabit örneklem genişliği testinde dikkate alınan kritik değere en yakın kritik değerlerin O'Brien & Fleming testi son analizinde dikkate alındığı, Pocock testi son analizinde dikkate alınan kritik değerlerin ise en yüksek farka sahip olduğu görülmektedir. Bu durum yapılması planlanan toplam analiz sayısına göre de değişmemektedir. $\alpha = 0,01$ I. tip hata düzeyinde, $K = 5$, $K = 10$, $K = 15$ ve $K = 20$ toplam analiz sayısı ile uygulanan O'Brien & Fleming ve Pocock testlerinin her bir ara analizde ve son analizde dikkate aldıkları kritik değerler sırasıyla Çizelge 4.5, Çizelge 4.6, Çizelge 4.7 ve Çizelge 4.8'de verilmiş, bu değerler Şekil 5.5, Şekil 5.6, Şekil 5.7 ve Şekil 5.8'de grafik üzerinde gösterilmiştir.

Eğer grup ardışık deneme düzeninin uygulandığı bir deneme boyunca yapılan ara analizlerde H_0 hipotezini reddedecek bir sonuca ulaşılamaz ve deneme son analize kadar devam ederse, sabit örneklem genişliği deneme düzeni kullanılan bir denemeye göre en az farkla karar verecek olan grup ardışık test tipi O'Brien & Fleming testidir. En uzak kritik değer ise Pocock testinde dikkate alınmaktadır.

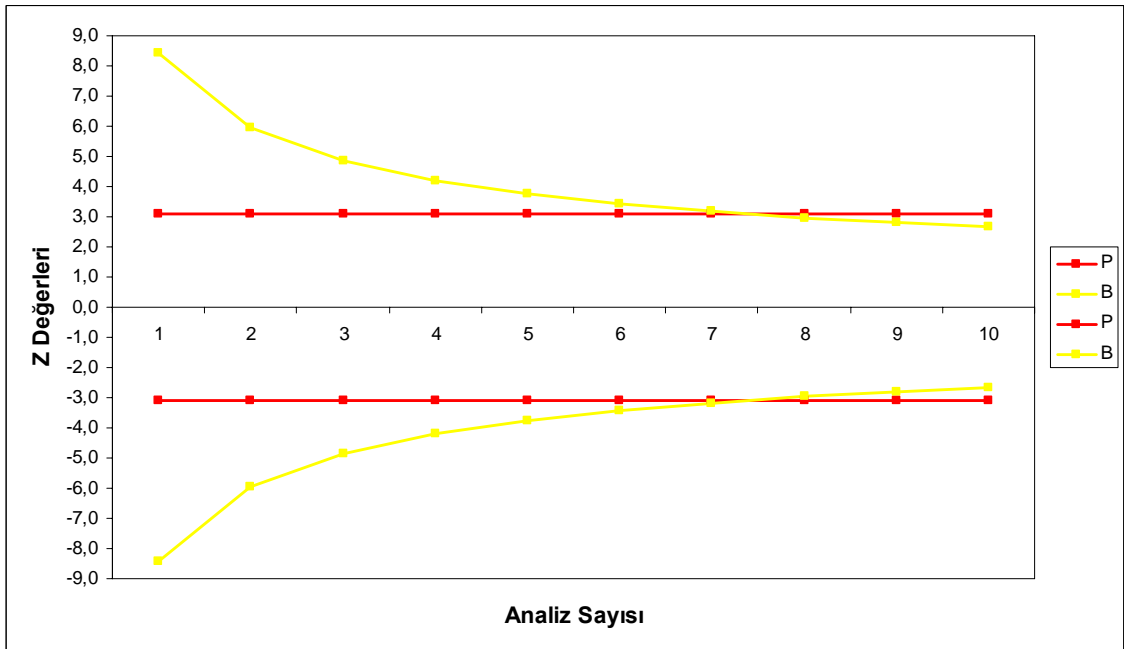
Grup ardışık deneme düzeninin uygulandığı bir deneme son analize kadar devam ederse O'Brien & Fleming testi, Pocock testine göre daha küçük etki büyüklüklerini tespit etme konusunda son analizde daha başarıyla ara analizlerde bu durum yine farklılık göstermekte ve sıralama her bir ara analizde farklı olarak ortaya çıkmaktadır.

Çizelge 4.5, Çizelge 4.6, Çizelge 4.7 ve Çizelge 4.8’de verilen, sırasıyla $K = 5$, $K = 10$, $K = 15$ ve $K = 20$ toplam analiz sayısı ve $\alpha = 0,01$ I. tip hata düzeyinde O’Brien & Fleming ve Pocock testlerinin her bir ara analiz ve son analizde dikkate aldıkları tüm kritik değerler incelendiğinde bu iki test tipinin birbirine göre durumları görülebilmektedir. Şekil 5.5, Şekil 5.6, Şekil 5.7 ve Şekil 5.8 incelendiğinde, ilk analizler için O’Brien & Fleming testinde dikkate alınan kritik değerler Pocock testine göre daha yüksek olduğundan, yine bu test tipinde H_0 hipotezini erken aşamalarda reddetmek oldukça zor görünmektedir. Yapılması planlanan toplam analiz sayısı (K) arttıkça ilk analizler için artan $\sqrt{(K/k)}$ değerine bağlı olarak $C_B(K, \alpha)/\sqrt{(K/k)}$ değerleri de artmakta ve ilk analizlerdeki farklılık K arttıkça daha da yüksek olmaktadır. Son (final) analize yaklaştıkça küçülen $\sqrt{(K/k)}$ değerine bağlı olarak küçülen $C_B(K, \alpha)/\sqrt{(K/k)}$ kritik değeri sona yakın ara analizlerde ve son analizde Pocock testinden daha küçük olmaktadır.

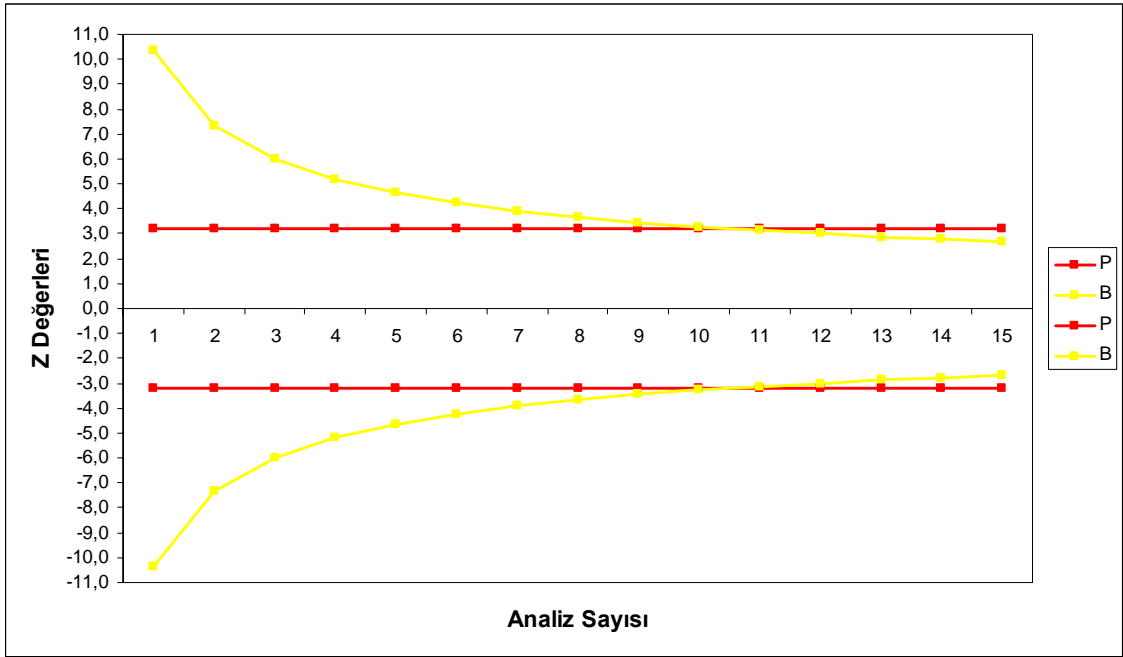
$\alpha = 0,01$ I. tip hata düzeyinde, $K = 5$ olduğunda ilk 3 ara analizde Pocock testi daha küçük etki büyüklüğünü tespit edebilecek kritik değerlere sahipken, 4. ara analiz ve son analizde O’Brien & Fleming testinin daha küçük etki büyüklüklerini tespit edebilecek duruma geldiği görülmektedir. $K = 10$ olduğunda bu değişim 8. ara analizden itibaren, $K = 15$ olduğunda 11. ara analizden itibaren ve $K = 20$ olduğunda ise 14. ara analizden itibaren ortaya çıkmaktadır. Dikkat edilirse $\alpha = 0,01$ I. tip hata düzeyinde, O’Brien & Fleming testi kritik değerlerinin Pocock testi kritik değerlerinden daha küçük olmaya başladığı ara analiz noktası $K = 5$ olduğunda $\alpha = 0,05$ I. tip hata düzeyindeki ile aynıdır (4. ara analiz). $K = 10$, $K = 15$ ve $K = 20$ olduğunda ise bu nokta $\alpha = 0,05$ I. tip hata düzeyindeki (sırasıyla 7., 10. ve 13. ara analizler) göre bir ara analiz daha sonra (sırasıyla 8., 11. ve 14. ara analizler) ortaya çıkmaktadır.



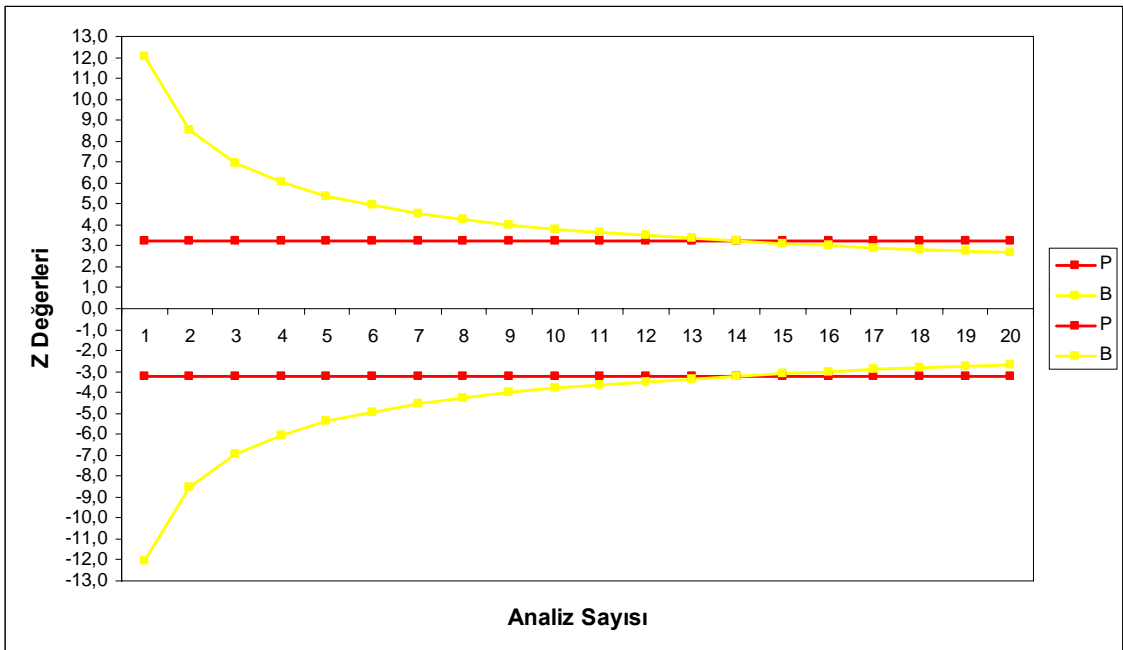
Şekil 5.5 $\alpha = 0,01$ ve $K = 5$ için Kritik Değerler



Şekil 5.6 $\alpha = 0,01$ ve $K = 10$ için Kritik Değerler



Şekil 5.7 $\alpha = 0,01$ ve $K = 15$ için Kritik Değerler



Şekil 5.8 $\alpha = 0,01$ ve $K = 20$ için Kritik Değerler

Grup ardışık testlerde her bir ara analizde, o anda mevcut olan verileri kullanarak yapılan ve o analiz için ardışık olmayan bir testin çift taraflı anlamlılık düzeyi, o teste ait I. tip hata olasılığından (α') küçükse H_0 hipotezi reddedilmektedir. Bu nedenle, bu test sabit bir anlamlılık düzeyi (α') ile tekrarlanan bir anlamlılık testidir. Buradaki çoklu inceleme nedeniyle, denemenin başında karar verilen toplam I. tip hata olasılığının (α) aşılmaması için, her bir ara analize ait ve toplam I. tip hata olasılığından (α) küçük bir anlamlılık düzeyinin (α'_k) belirlenmesi zorunludur (1, 2, 3, 5, 36). Örneğin; Pocock testi için $K = 5$ ve $\alpha = 0,05$ toplam I. tip hata düzeyinde, her bir ara analizde dikkate alınan $C_p(K, \alpha) = 2,413$ değerine karşılık gelen anlamlılık düzeyi $\alpha'_k = 0,0158$ 'dır ve her bir ara analiz için sabittir. O'Brien & Fleming testinde ise, k arttıkça $C_p(K, \alpha) / \sqrt{(K/k)}$ değeri azalacağından α'_k artacaktır. Yine $K = 5$ ve $\alpha = 0,05$ toplam I. tip hata düzeyinde, her bir ara analizde anlamlılık düzeyleri $\alpha'_1 = 0,000005$, $\alpha'_2 = 0,0013$, $\alpha'_3 = 0,0084$, $\alpha'_4 = 0,0225$ ve $\alpha'_5 = 0,0413$ şeklinde olacaktır (3).

Bir grup ardışık deneme düzeni kullanmanın başlıca nedeni örneklem genişliğini azaltmaktır. Örneklem genişliği dağılımı ise esas olarak etki büyüklüğüne (d) dayanmaktadır ve karşılaştırmalar yapılırken etki büyüklüğünün (d) bir aralığı için örneklem genişlikleri dikkate alınmalıdır. Maksimum örneklem genişliği en kötü durumda ne olabileceğinin bir belirtisidir (3).

Bir denemede kullanılacak olan örneklem genişliği belirlenirken, I. tip hata düzeyi (α), test gücü ($1 - \beta$), etki büyüklüğü (d) ve varyansın (σ^2) farklı değerleri için gerek duyulan örneklem genişliği de farklılık göstermektedir. Bu dört parametreden I. tip hata düzeyi (α) ve etki büyüklüğünün (d) gerek duyulan örneklem genişliğine etkisi ters yöndeyken test gücü ($1 - \beta$) ve varyansın (σ^2) etkisi ise buna aksi yöndedir. Yani artan I. tip hata düzeyi (α) ve etki büyüklüğü (d) için gerek duyulan örneklem genişliği azalırken, artan test gücü ($1 - \beta$) ve varyans (σ^2) değerleri için gerek duyulan örneklem genişliği de artmaktadır (16, 35).

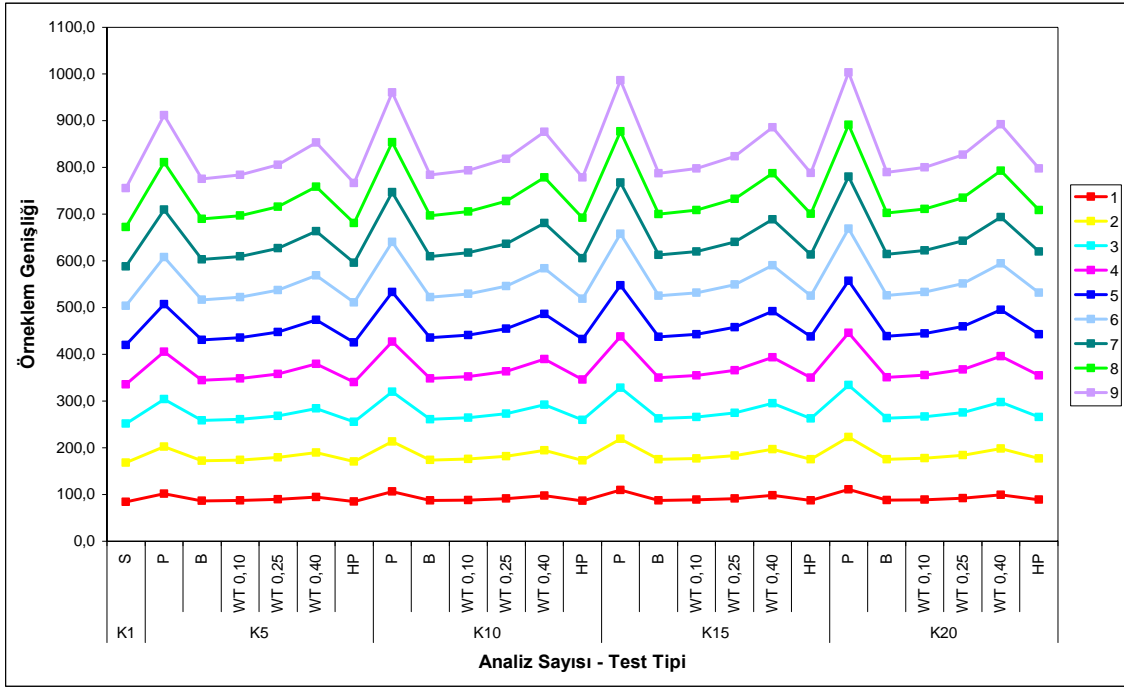
Grup ardışık test tiplerinde kullanılabilen maksimum örneklem genişliğinin birbirleriyle ve sabit örneklem genişliği testinde gerek duyulan örneklem genişliği ile karşılaştırılması amacıyla farklı I. tip hata düzeyi (α), test gücü ($1-\beta$), etki büyüklüğü (d) ve varyans (σ^2) seçimleri için dikkate alınan kombinasyonlarda yukarıda verilen literatür bilgisine aykırı bir durum gözlenmemiştir. Çizelge 4.9'dan Çizelge 4.32'ye kadar olan yirmi dört çizelgede dikkate alınan tüm kombinasyonlar için, grup ardışık test tiplerinde gerek duyulabilecek maksimum örneklem genişlikleri tek tek verilmiş, Şekil 5.9'dan Şekil 5.32'ye kadar olan şekillerde ise bu değerler grafik üzerinde gösterilmiştir. $K=1$ için hesaplanan örneklem genişlikleri, sabit örneklem genişliği deneme düzenlerinde gerek duyulan örneklem genişliğine eşittir. Bölüm 2.6.3 ve alt bölümlerinde verilen bilgiler dikkate alındığında, K arttıkça grup ardışık test tiplerinde maksimum örneklem genişliğinin de artacağı açıktır. Ancak grup ardışık deneme düzenlerindeki esas amaç, yapılan ara analizler sayesinde daha erken durma olasılığı ve örneklem genişliğinde azalma avantajını değerlendirmektir (31, 32, 40).

Çizelgeler ve şekiller incelendiğinde, aynı I. tip hata düzeyi (α), test gücü ($1-\beta$), etki büyüklüğü (d) ve varyans (σ^2) koşulları altında, grup ardışık test tiplerinin maksimum örneklem genişliklerinin yapılması planlanan toplam analiz sayısına (K) göre de farklılık gösterdiği görülmektedir. Genel olarak belirli bir I. tip hata düzeyi (α), test gücü ($1-\beta$), etki büyüklüğü (d) ve varyans (σ^2) seçiminde sabit örneklem genişliği testinde gerek duyulan örneklem genişliğine en yakın maksimum örneklem genişliği gereksinimi O'Brien & Fleming ve Haybittle-Peto testlerinde ortaya çıkmaktadır. Pocock testi ise en yüksek maksimum örneklem genişliğine ihtiyaç duyan grup ardışık test tipidir. Bu durum, I. tip hata düzeyi (α) ve test gücüne ($1-\beta$) göre farklılık göstermemektedir. Etki büyüklüğü (d), varyans (σ^2) ve toplam analiz sayısı (K) bakımından ise sonuçlar farklılık göstermektedir.

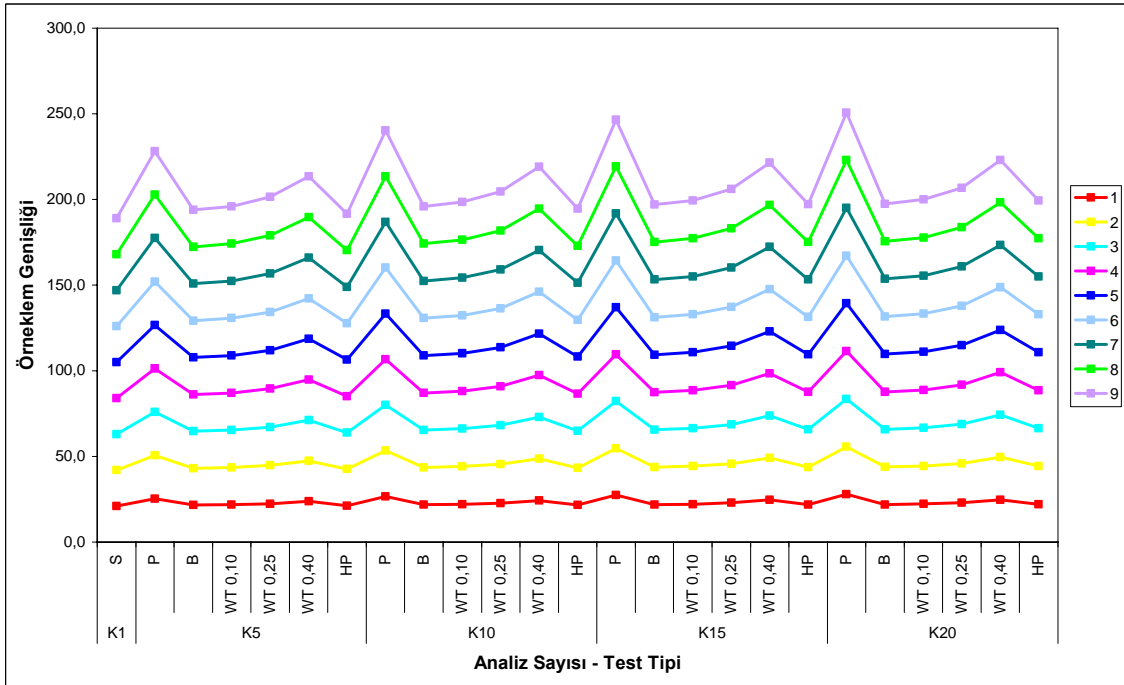
Örneğin; $\alpha=0,05$ I. tip hata olasılığı ve $(1-\beta)=0,90$ test gücü koşullarında, $0,5 \leq d \leq 1,5$ için, $K=10$ analize kadar Haybittle-Peto testi, O'Brien & Fleming testinden daha küçük maksimum örneklem genişliği gereksinimine sahipken, $K > 10$ için bu durum tersine dönmekte ve O'Brien & Fleming testi en küçük maksimum

örneklem genişliği gereksinimine sahip olmaktadır. Etki büyüklüğü arttıkça bu yer değiştirme de daha büyük K sayıları için olmakta, $1,5 > d \geq 2,5$ için $K = 15$ analize kadar Haybittle-Peto testi, en küçük maksimum örneklem genişliğine sahipken, $K > 15$ için O'Brien & Fleming testi en küçük maksimum örneklem genişliğine sahip olmaktadır. $d \geq 3,0$ için ise O'Brien & Fleming testinde gerek duyulan maksimum örneklem genişliği, $K \geq 20$ için en küçük olabilmektedir. Ancak burada dikkat edilmesi gereken nokta artan etki büyüklüğüne bağlı olarak, gerek duyulan örneklem genişliklerinin de azaldığı ve yüksek etki büyüklüğü değerleri için bu karşılaştırmaların teorik düzeyde yapılmış olduğudur. $d = 3,0$ etki büyüklüğü için hesaplanan örneklem genişliği değerleri 1,7 düzeylerine kadar düşebilmektedir. Uygulamada, örneklem genişliği hesaplandıktan sonra elde edilen değer bir üst tam sayıya yuvarlanarak kullanıldığından grup ardışık test tipleri arasındaki bu farklılıklar ortadan kalkacaktır. Burada ihmal edilmemesi gereken bir diğer konu ise bu durumun, deneme boyunca yapılan ara analizlerde H_0 hipotezini reddedecek bir sonuca ulaşılamayıp son analize kadar gelindiğinde ortaya çıkacak olmasıdır. Etki büyüklüğünün (d) yüksek, veriler arasındaki varyasyonun da düşük olması durumunda ilk ara analizlerde denemenin durma olasılığı çok yüksek olduğundan, bütün grup ardışık deneme düzenlerinin sabit örneklem genişliği deneme düzenlerine göre daha az deneğe gerek duyacağı açıktır.

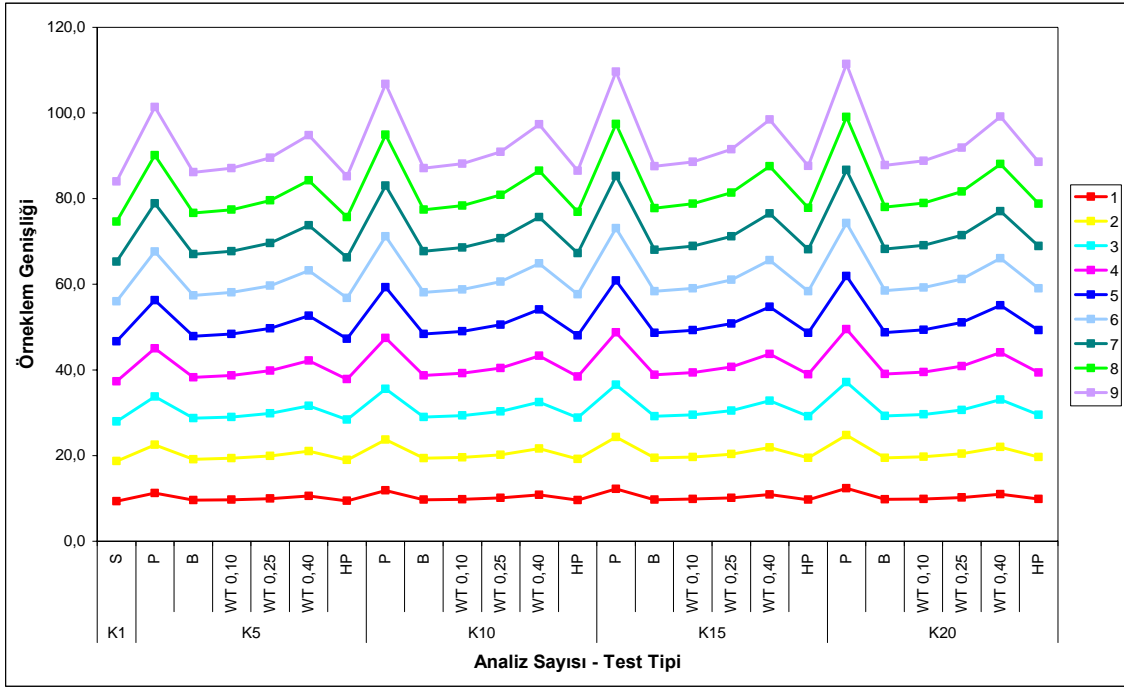
Yapılan tüm kombinasyonlarda elde edilen sonuçlar genel olarak incelendiğinde de benzer sonuçlar ortaya çıkmaktadır. Haybittle-Peto ve O'Brien & Fleming testleri diğer test tiplerine göre belirgin şekilde daha küçük maksimum örneklem genişliklerine sahiptir. Haybittle-Peto ve O'Brien & Fleming testlerinin arasındaki farklılık artan etki büyüklüğü koşullarında önemsenmeyecek kadar azalmaktadır. Tüm kombinasyonlarda en büyük maksimum örneklem genişliği gereksinimi ise Pocock testinde görülmektedir. Wang & Tsatis testi daima bu test tiplerinin arasında kalan bir maksimum örneklem genişliğine sahip olmaktadır. Farklı Δ seçimleri Wang & Tsatis testinin durumunu değiştirmemekte, artan Δ değerleri maksimum örneklem genişliğinde artmalara sebep olurken daha küçük Δ değerleri de daha az maksimum örneklem genişliği gereksinimleri ile sonuçlanmaktadır.



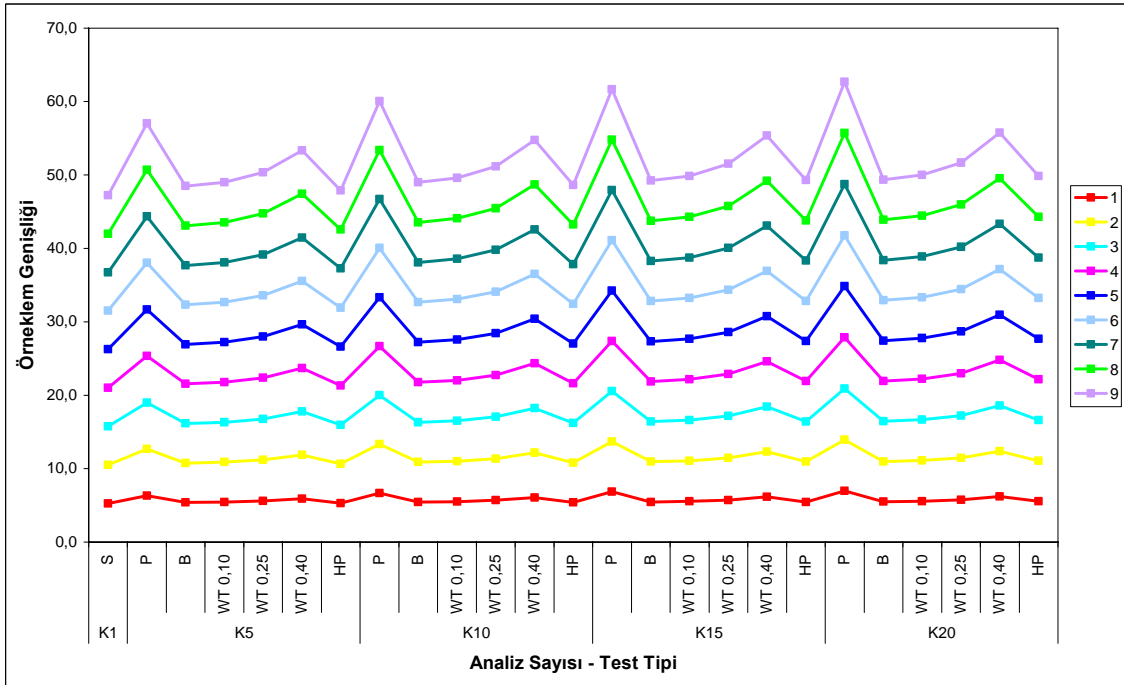
Şekil 5.9 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 0,5$ için Örneklem Genişlikleri



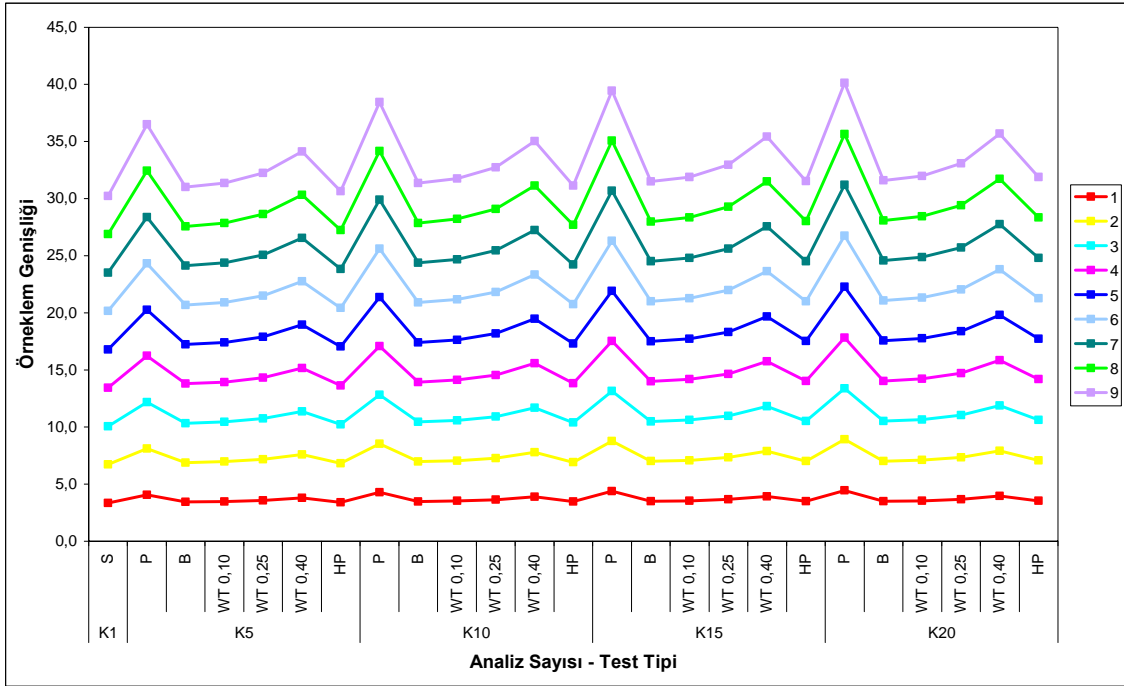
Şekil 5.10 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 1,0$ için Örneklem Genişlikleri



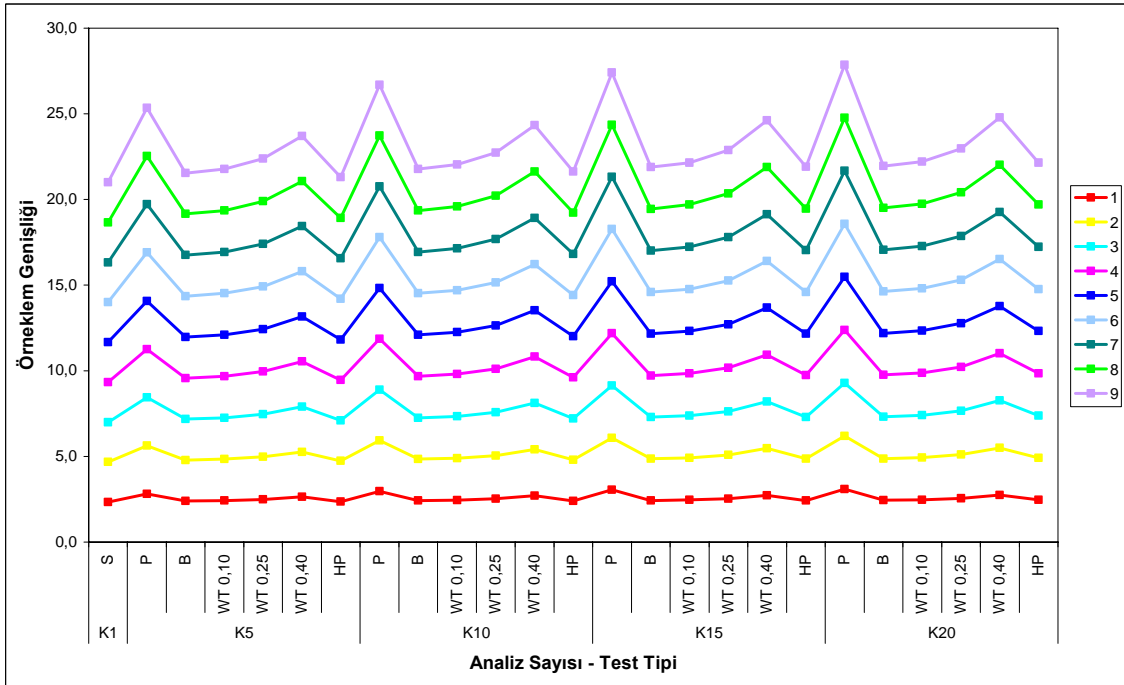
Şekil 5.11 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 1,5$ için Örneklem Genişlikleri



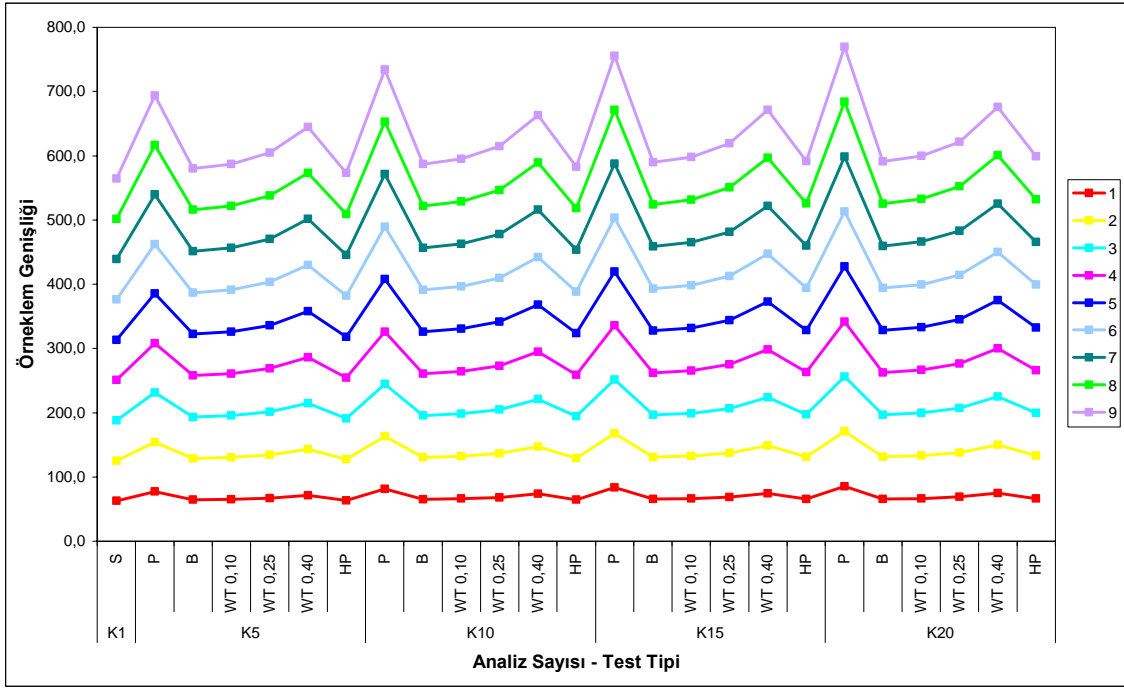
Şekil 5.12 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 2,0$ için Örneklem Genişlikleri



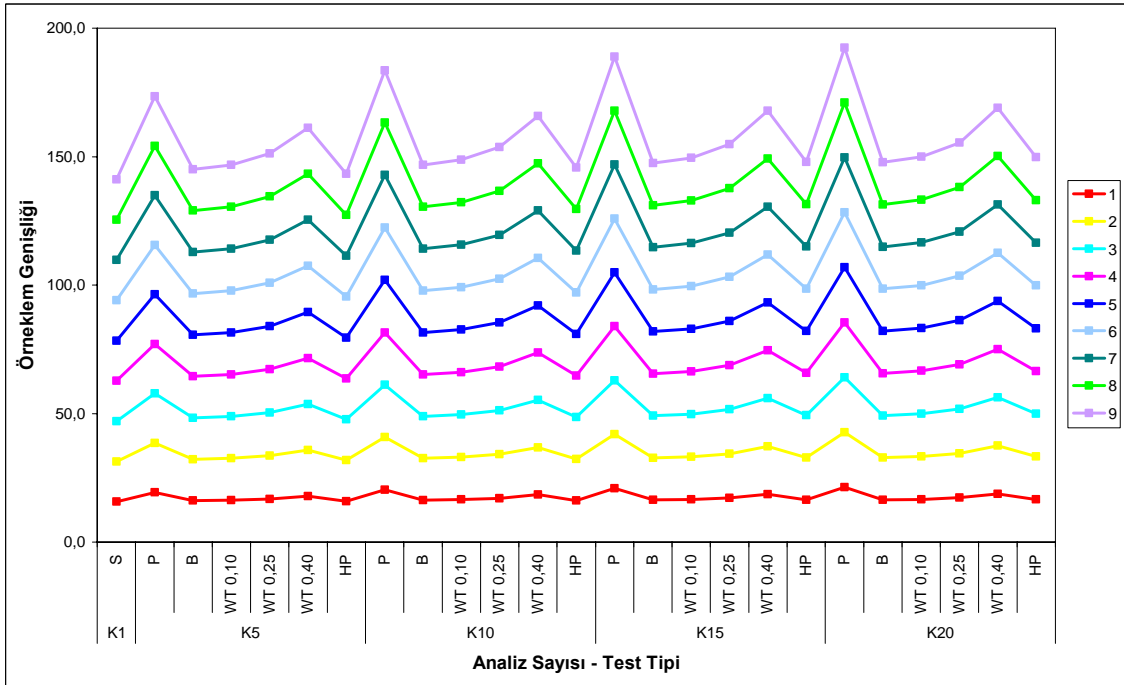
Şekil 5.13 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 2,5$ için Örneklem Genişlikleri



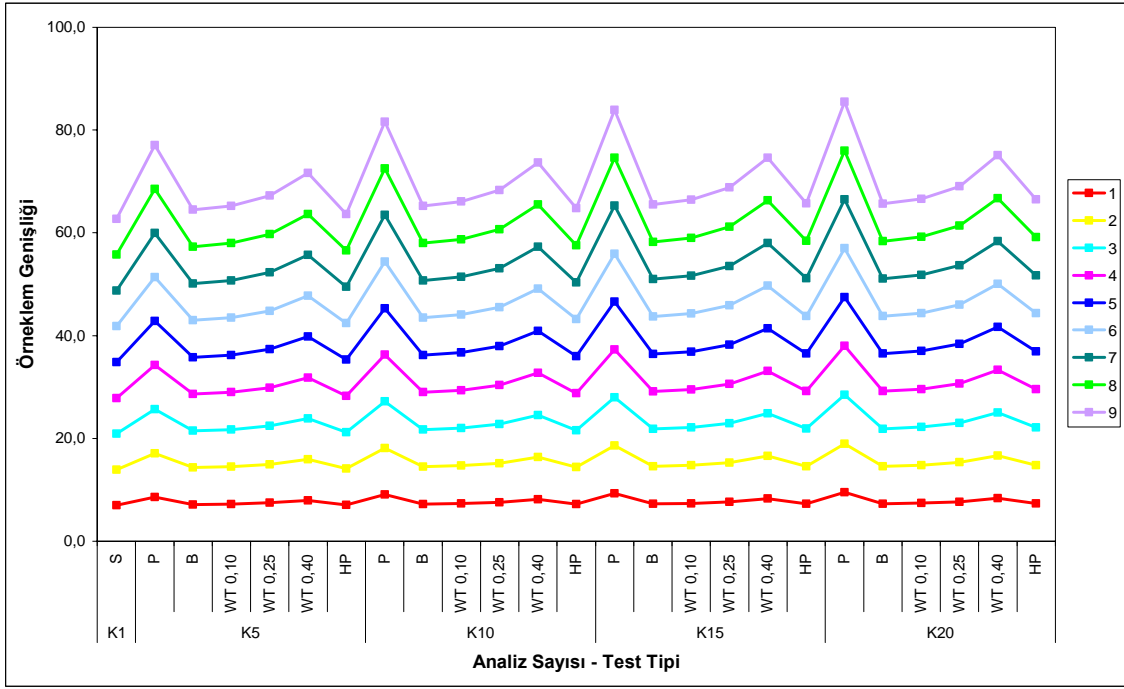
Şekil 5.14 $\alpha = 0,05$, $(1 - \beta) = 0,90$ ve $d = 3,0$ için Örneklem Genişlikleri



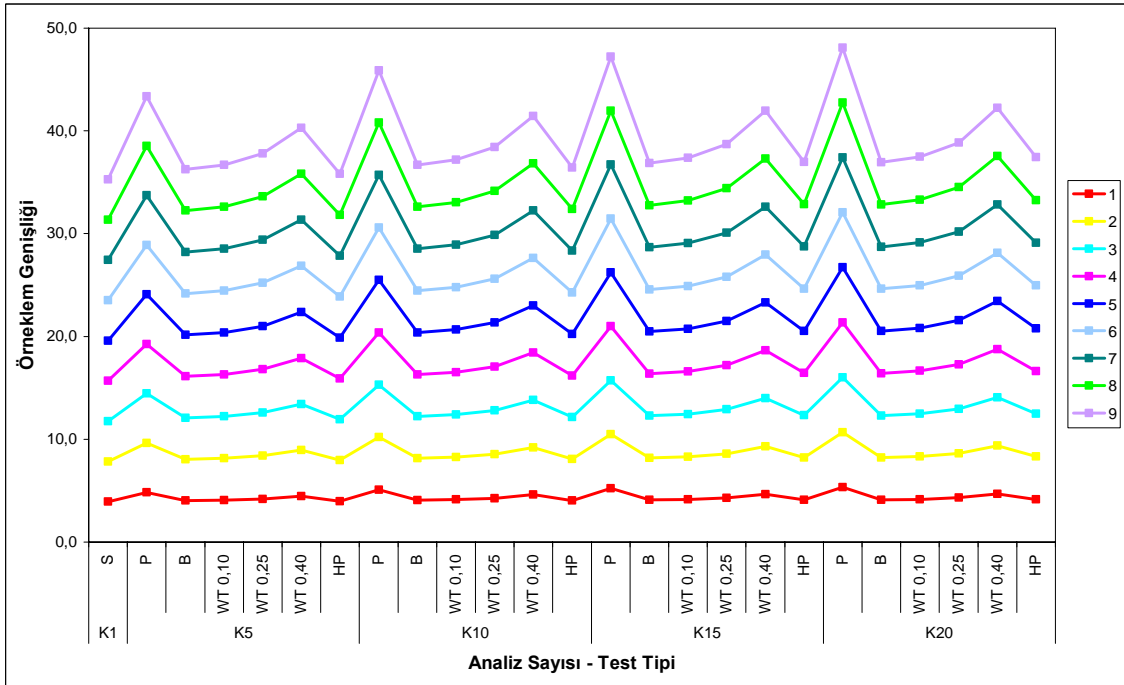
Şekil 5.15 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 0,5$ için Örneklem Genişlikleri



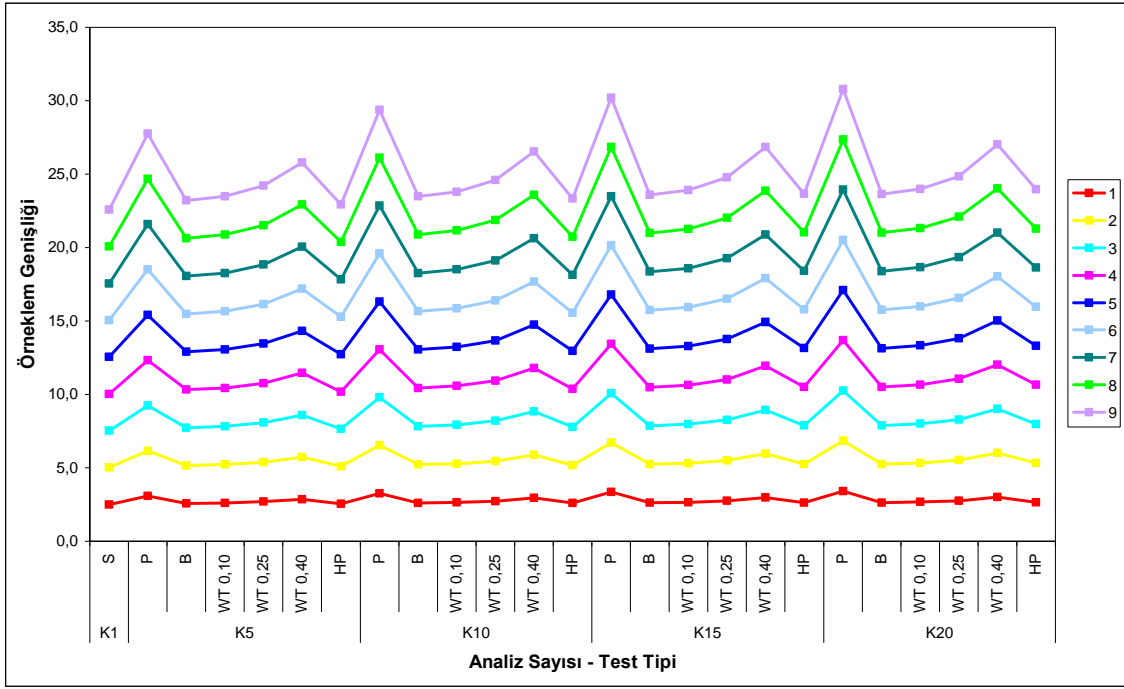
Şekil 5.16 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 1,0$ için Örneklem Genişlikleri



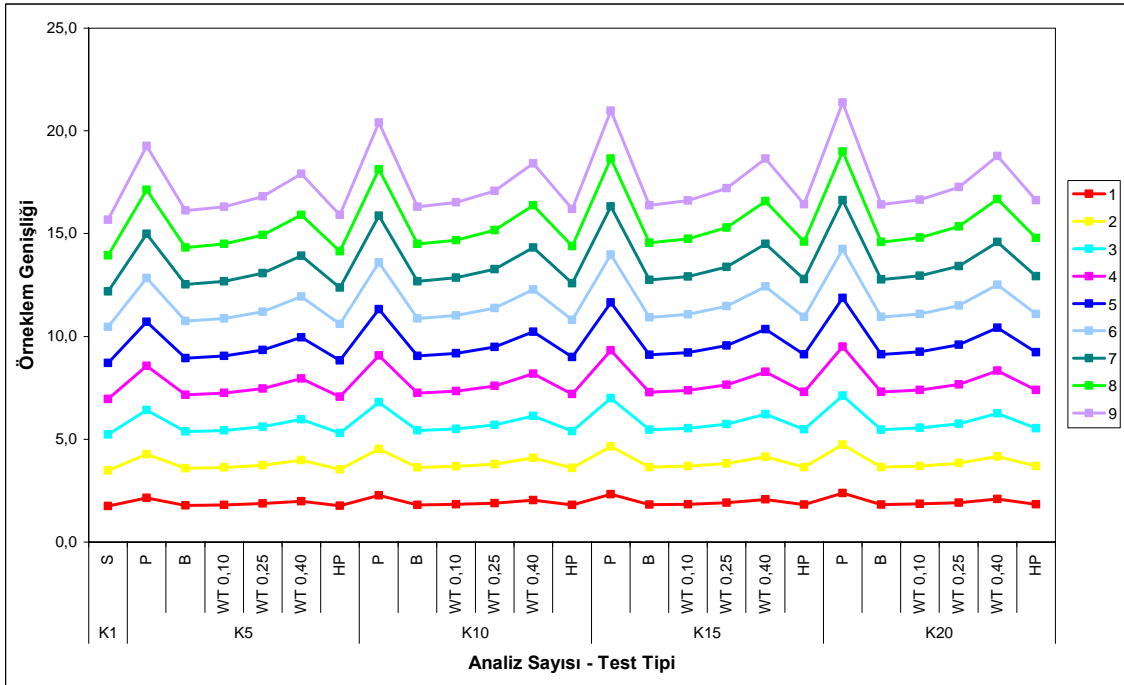
Şekil 5.17 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 1,5$ için Örneklem Genişlikleri



Şekil 5.18 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 2,0$ için Örneklem Genişlikleri



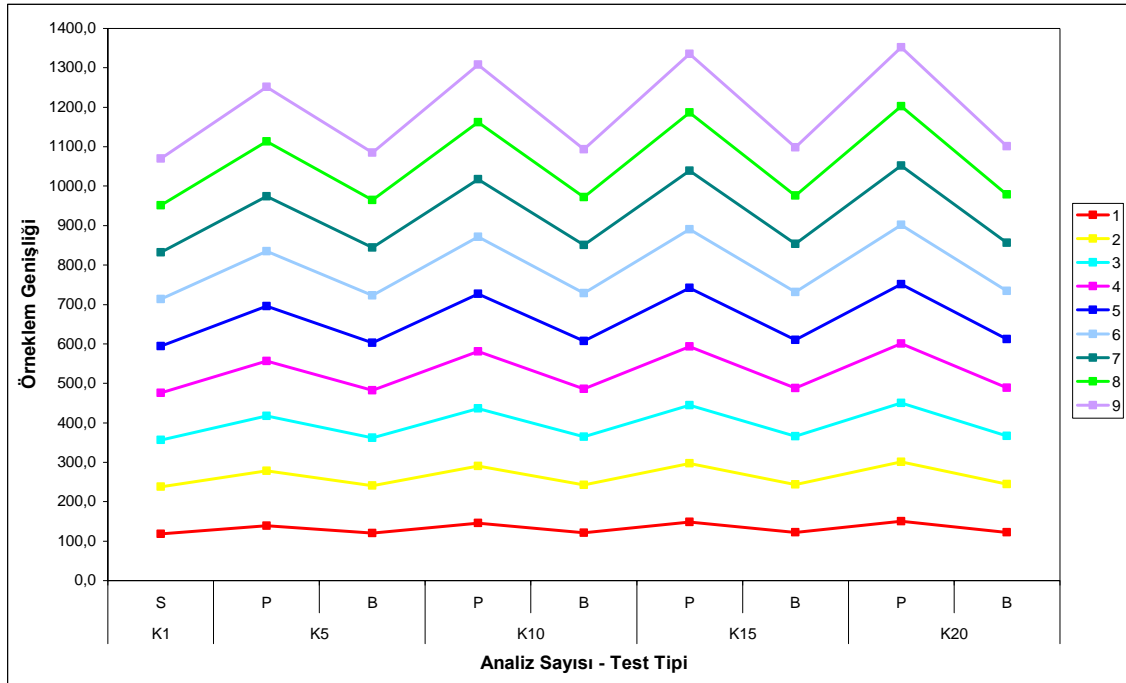
Şekil 5.19 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 2,5$ için Örneklem Genişlikleri



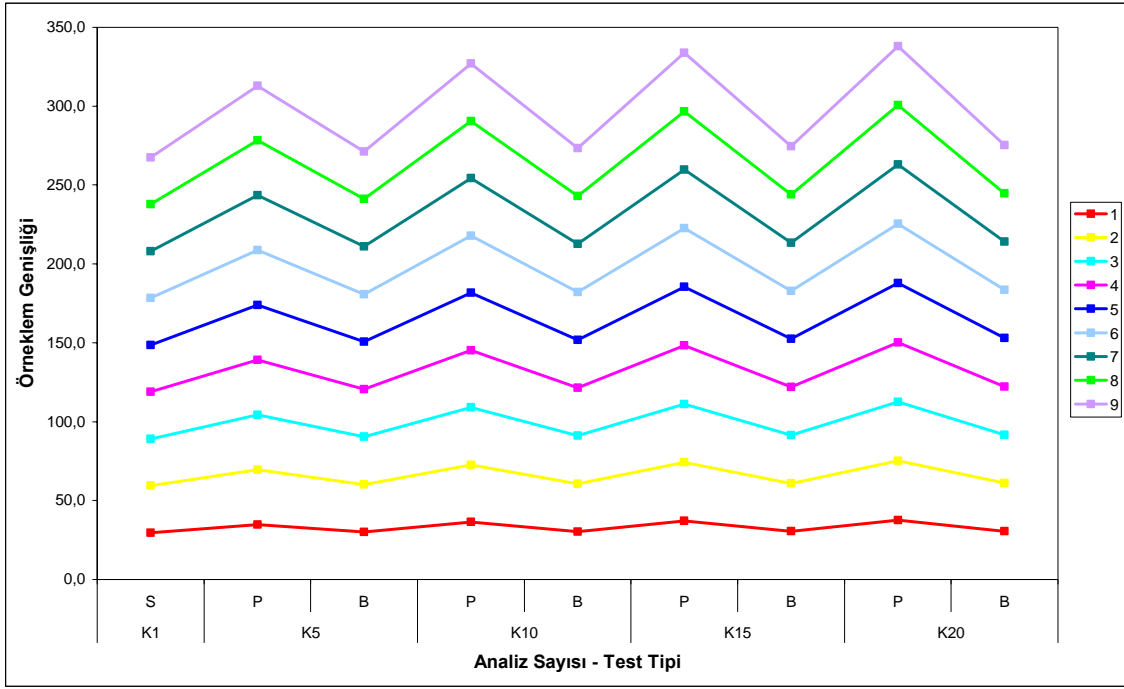
Şekil 5.20 $\alpha = 0,05$, $(1 - \beta) = 0,80$ ve $d = 3,0$ için Örneklem Genişlikleri

Bölüm 2.6.3.4'te açıklandığı gibi $\alpha = 0,05$ düzeyinden daha düşük bir I. tip hata olasılığı sağlanamadığından Haybittle-Peto testi için maksimum örneklem genişliği sadece $\alpha = 0,05$ I. tip hata düzeyinde hesaplanabilmektedir. Wang & Tsatis testleri ise farklı Δ değerleri için Pocock ve O'Brien & Fleming testlerinin arasında yer aldığından $\alpha = 0,01$ I. tip hata düzeyinde sadece Pocock ve O'Brien & Fleming testleri birbirleriyle ve sabit örneklem genişliği deneme düzeni testiyle karşılaştırılmıştır.

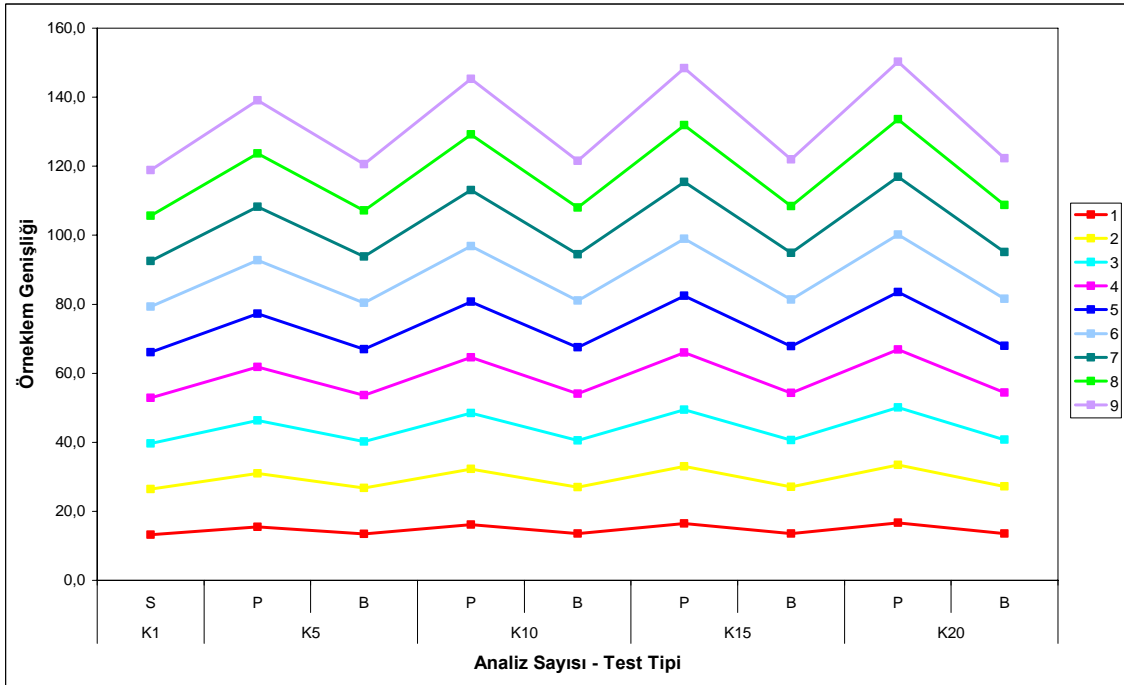
$\alpha = 0,01$ I. tip hata düzeyinde yapılan tüm kombinasyonlarda da sonuçlar farklılık göstermemiş, aynı test gücü ($1 - \beta$), etki büyüklüğü (d) ve varyans (σ^2) koşulları altında O'Brien & Fleming testinde maksimum örneklem genişliği gereksiniminin Pocock testinden daima küçük olduğu sonucuna ulaşılmıştır. Yapılması planlanan toplam analiz sayısına (K) göre maksimum örneklem genişliği gereksinimlerinde ortaya çıkan değişim O'Brien & Fleming ve Pocock testleri arasında herhangi bir farklılığa neden olmamıştır.



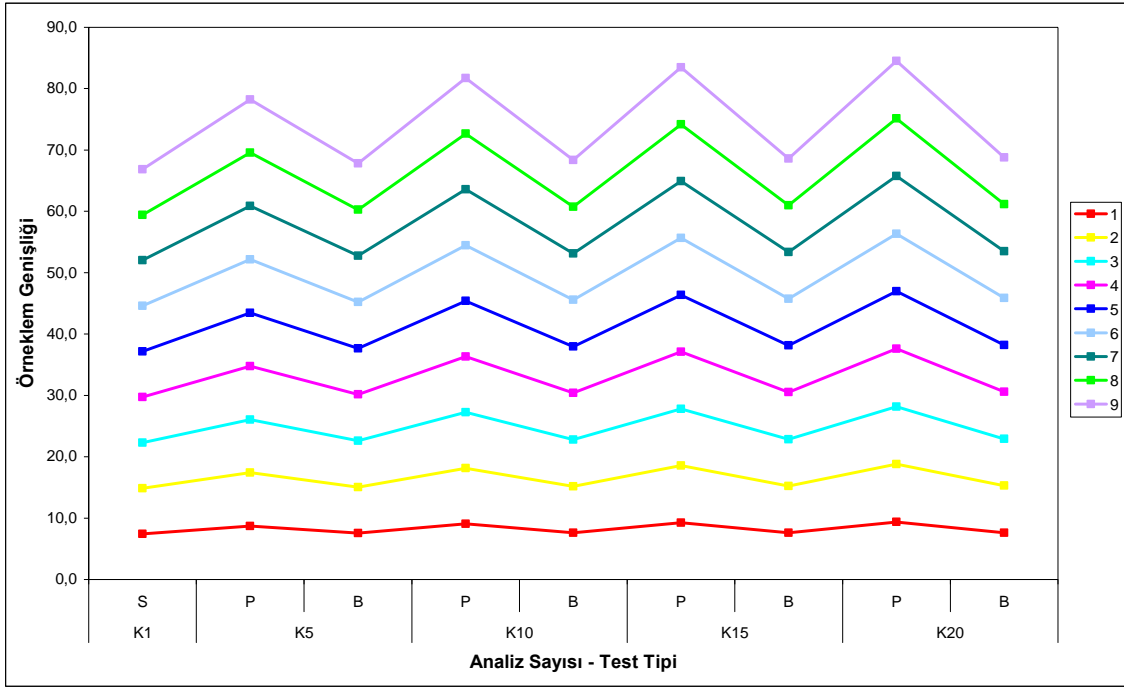
Şekil 5.21 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 0,5$ için Örneklem Genişlikleri



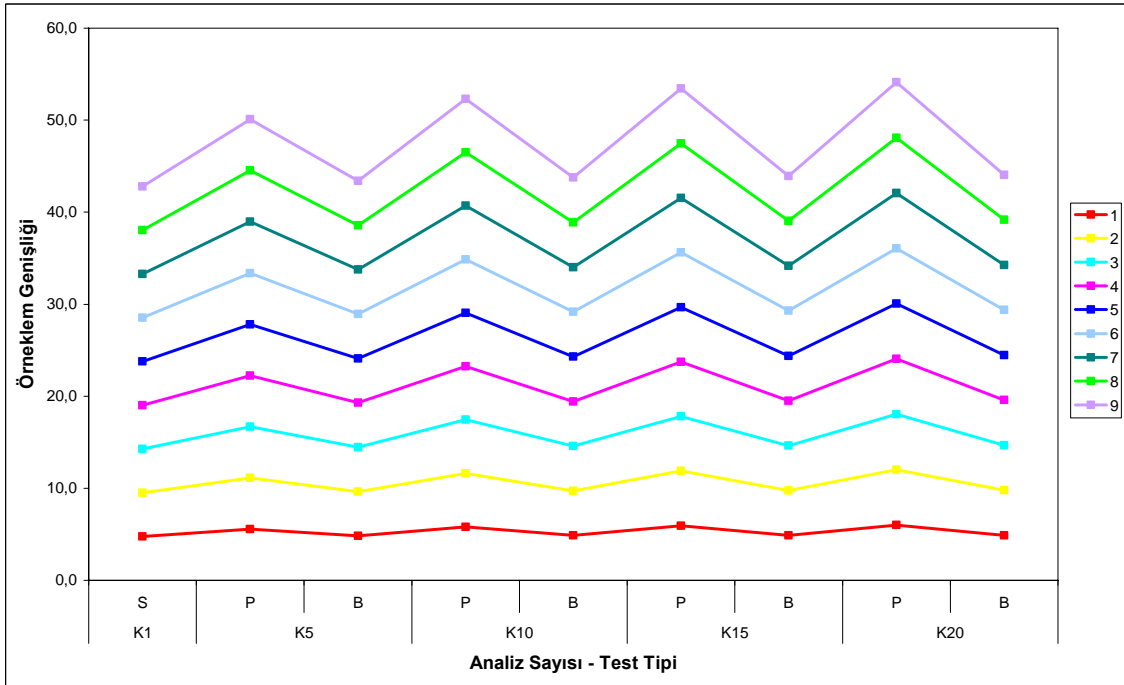
Şekil 5.22 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 1,0$ için Örneklem Genişlikleri



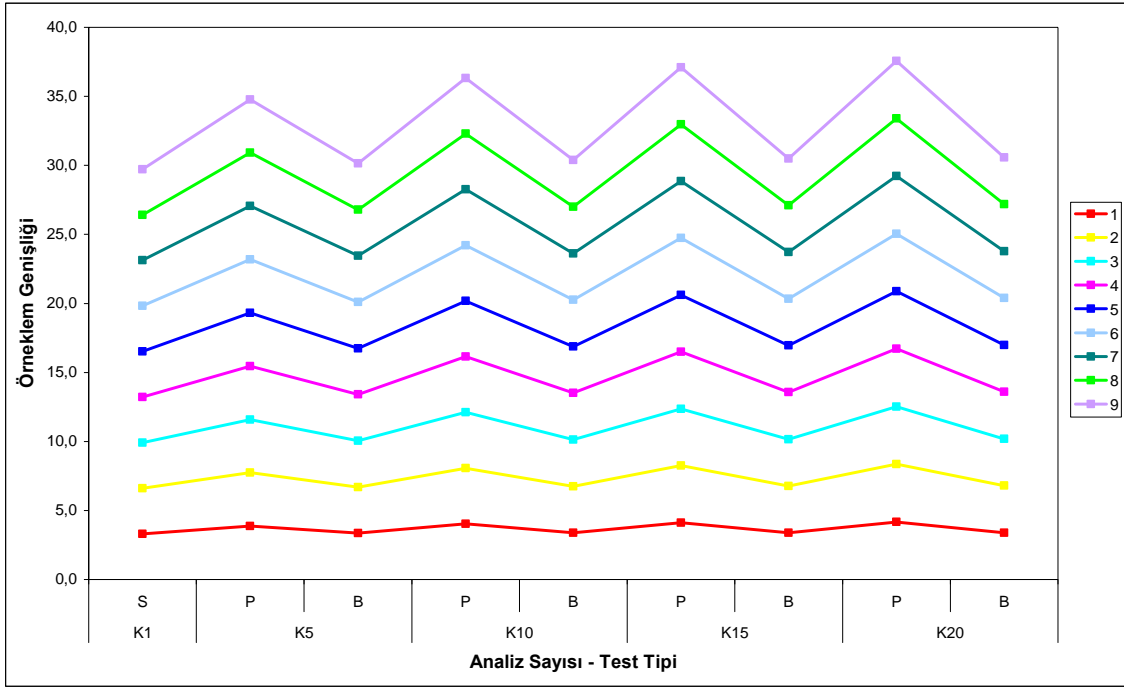
Şekil 5.23 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 1,5$ için Örneklem Genişlikleri



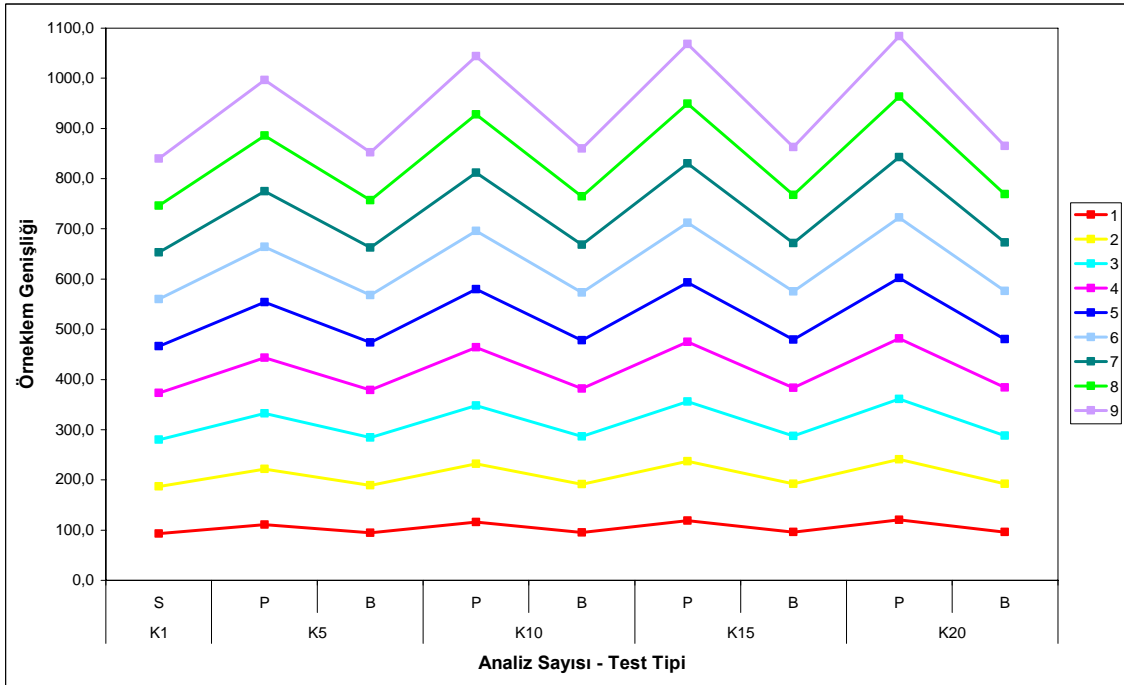
Şekil 5.24 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 2,0$ için Örneklem Genişlikleri



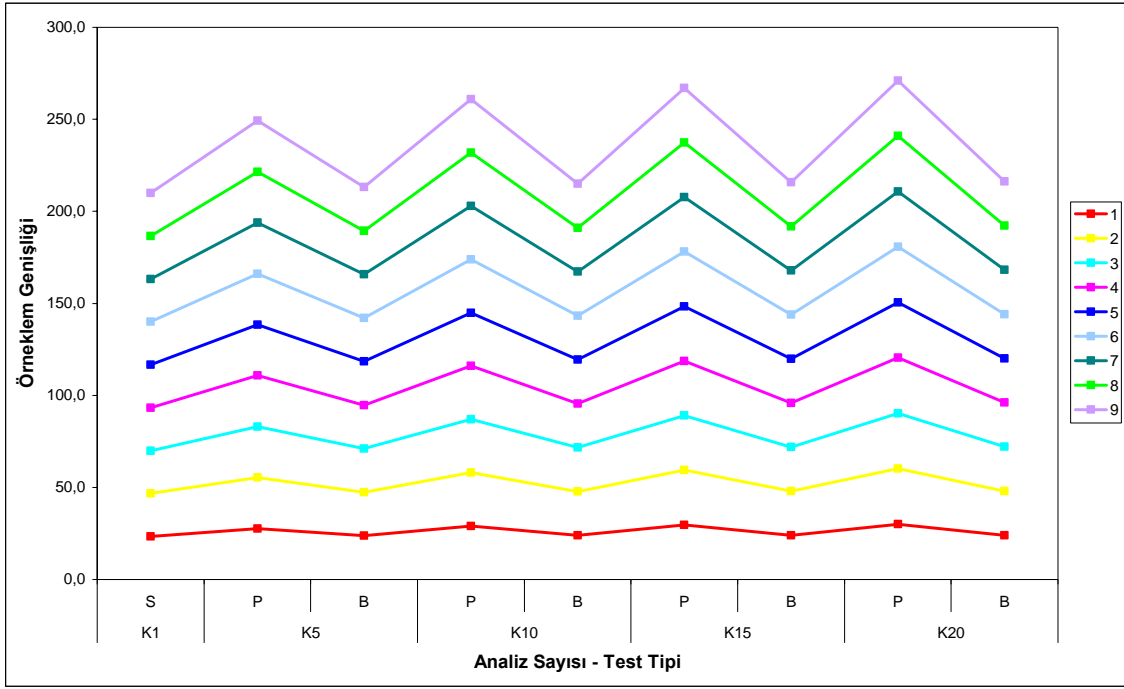
Şekil 5.25 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 2,5$ için Örneklem Genişlikleri



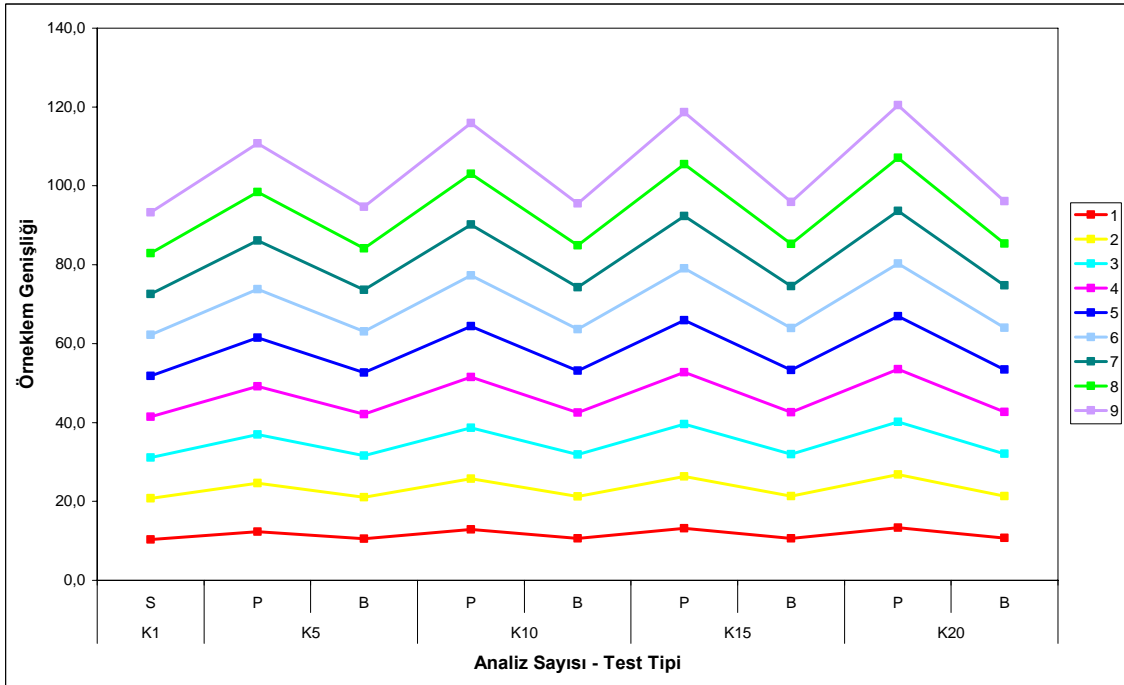
Şekil 5.26 $\alpha = 0,01$, $(1 - \beta) = 0,90$ ve $d = 3,0$ için Örneklem Genişlikleri



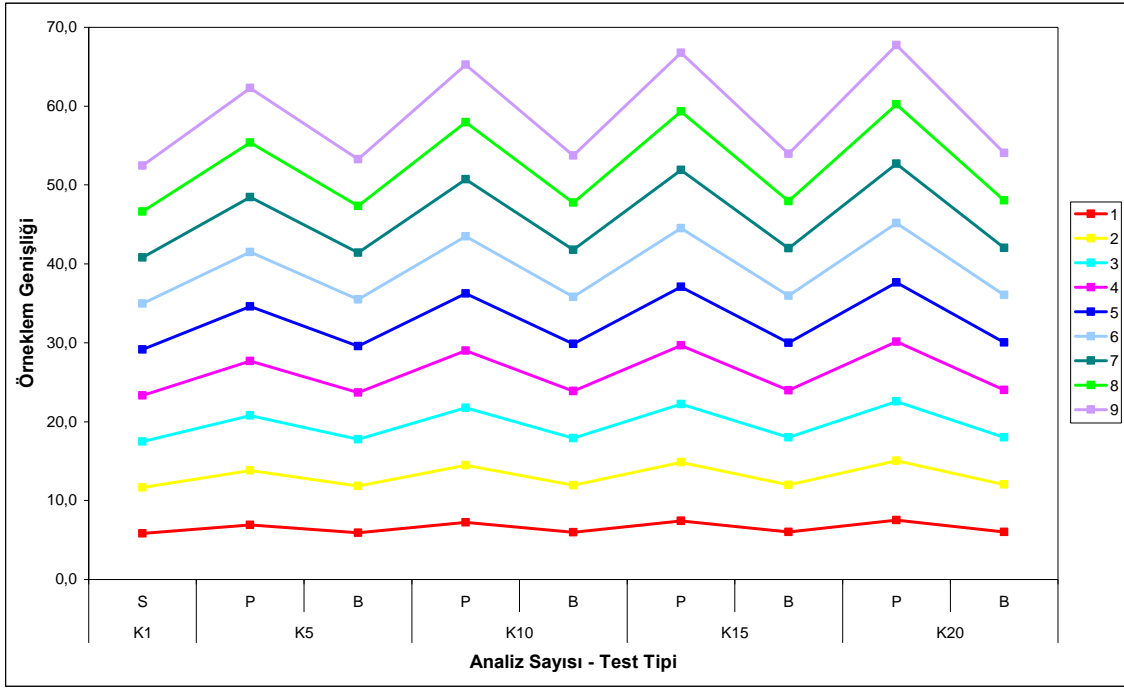
Şekil 5.27 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 0,5$ için Örneklem Genişlikleri



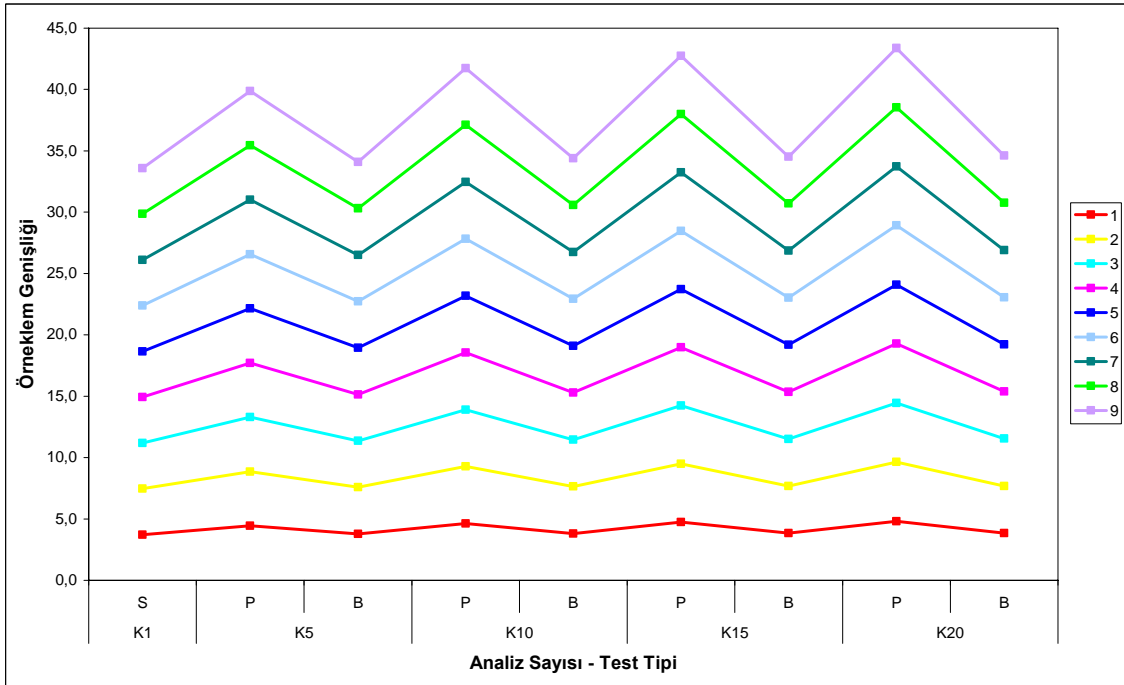
Şekil 5.28 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 1,0$ için Örneklem Genişlikleri



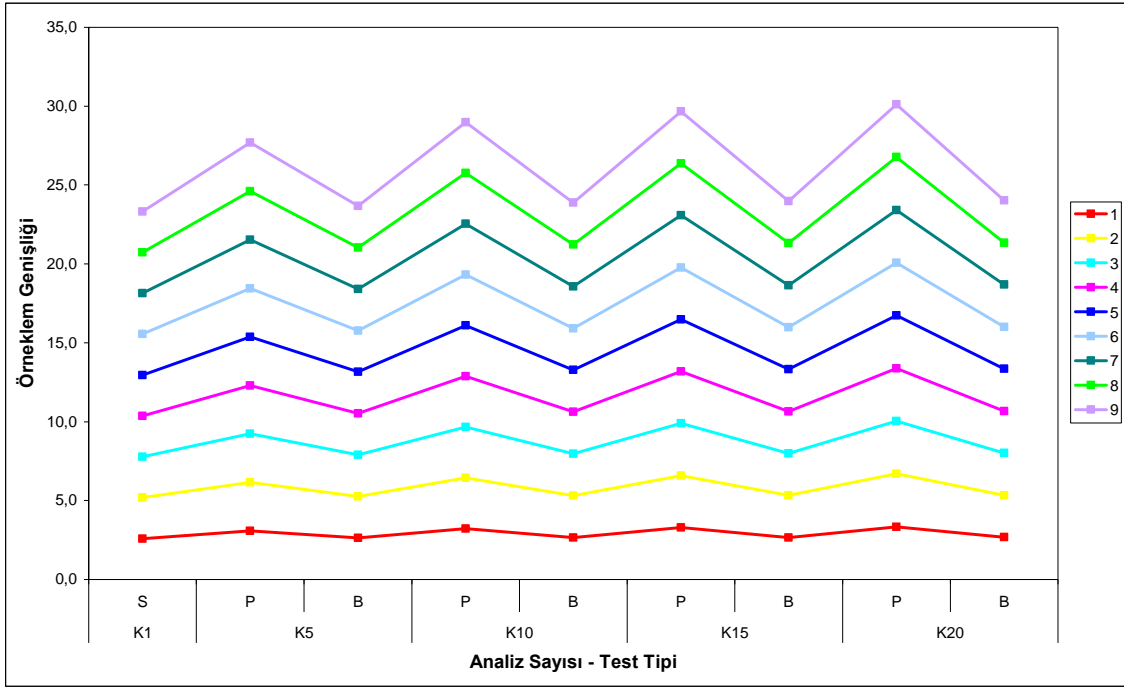
Şekil 5.29 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 1,5$ için Örneklem Genişlikleri



Şekil 5.30 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 2,0$ için Örneklem Genişlikleri



Şekil 5.31 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 2,5$ için Örneklem Genişlikleri



Şekil 5.32 $\alpha = 0,01$, $(1 - \beta) = 0,80$ ve $d = 3,0$ için Örneklem Genişlikleri

O'Brien & Fleming testinin, Pocock testinden daha küçük maksimum örneklem genişliği gerektirmesinin nedeni, grup ardışık testlerde test istatistiğinin hesaplandığı formülden anlaşılabilir. Eşitlik (2.4) incelendiğinde, etki büyüklüğünün, test istatistiğinin beklenen değeri ($E(Z_k)$) üzerine etkisinin analiz sayısı ile (k) birlikte arttığı görülmektedir, bu nedenle testin gücü esas olarak sonraki analizlerden kazanılmaktadır (3, 19). O'Brien & Fleming testinin kritik değer sınırları ilk ara analizlerde daha geniştir, fakat aynı örneklem genişliği ile Pocock testinin kritik değerler sınırlarından daha fazla güç kazanabildiği son birkaç ara analizde ve son analizde daha dardır. Bu nedenle, belirli bir test gücünde, Pocock testi çok büyük maksimum örneklem genişliği gereksinimine sahipken O'Brien & Fleming testi daha küçük maksimum örneklem genişliğine gerektirmektedir (3, 19).

Buradaki karşılaştırmalar yapılırken her bir kombinasyon için ayrı ayrı belirlenen I. tip hata olasılıkları (α), test gücü ($1 - \beta$), etki büyüklükleri (d) ve varyans (σ^2) değerleri için hesaplamalardan elde edilen örneklem genişlikleriyle çalışmanın

mümkün olduğu varsayıldı. Uygulamada, herhangi bir test için maksimum süre ve maksimum örneklem genişliği belirlenirken zaman kısıtlamaları veya kaynakların sınırlı olması gibi nedenlerle, bazı çalışmalarda bu şartlar tam olarak gerçekleştirilemeyebilir. Çoğu klinik denemelerde, minimum güç gereksinimini garanti etmek için yeterli büyüklükte bir örneklem genişliğine ulaşmanın zor olduğu, uygulamada çoğu zaman mevcut olan deneklerle çalışıldığı da göz ardı edilmemelidir. Böyle durumlarda ardışık deneme düzenleri kullanılacaksa, beklenen örneklem genişliğini azaltmak için test gücünden taviz vermek gerekebilmektedir (3).

6. SONUÇLAR VE ÖNERİLER

Genel olarak grup ardışık test tiplerinde maksimum örneklem genişlikleri, sabit örneklem genişliği testinde gerek duyulan örneklem genişliğinden daha fazladır. Ancak bu durum, H_0 hipotezini kabul etmek için grup ardışık testlerin maksimum örneklem genişliğine kadar devam etmek zorunluluğu olduğundan beklenen bir durumdur. Bununla birlikte, muamele farklılığı büyük olduğunda denemeyi erken sonlandırmak için ilk ara analizlerde verilecek erken bir karar, o aşamadan sonraki analiz grubundaki deneklerin etkin olmayan bir muameleye maruz kalmasını önleyecektir. Etki büyüklüğünü sıfırdan uzaklaştıkça grup ardışık testler, sabit örneklem genişliği testine göre kullanılan denek sayısında dikkate değer azalmalar sağlamaktadır.

Aynı şartlar altında grup ardışık test tiplerinin birbirine göre durumları incelendiğinde en fazla maksimum örneklem genişliği gereksiniminin Pocock testinde ortaya çıktığı görülmektedir. Uygulamada, Pocock testi örneklem genişliğindeki bu büyük gereksinimleri elde etmek bazen mümkün olmayabilir ve daha yüksek maksimum örneklem genişliği ciddi bir dezavantaj olarak ortaya çıkacaktır. Haybittle-Peto testinin en önemli dezavantajı $\alpha = 0,01$ I. tip hata düzeyinde kullanılamıyor olmasıdır. Wang & Tsatis testi ise $\Delta = 0,50$ seçiminde Pocock testine ve $\Delta = 0,00$ seçiminde O'Brien & Fleming testine eşit olmak üzere, değişen Δ değerleri için, arada kalan sonuçları vermektedir. Bu nedenle daha çok Pocock testi ve O'Brien & Fleming testinin birbiriyle ve sabit örneklem genişliği deneme düzen testiyle karşılaştırılması, grup ardışık test tiplerinin birbirine göre avantaj ve dezavantajlarının görülmesini sağlayacaktır.

Örneğin; $\alpha = 0,01$ I. tip hata olasılığı, $(1 - \beta = 0,90)$ test gücü, $d = 1$ etki büyüklüğü ve $\sigma^2 = 3$ varyans değerleri için hesaplanacak örneklem genişlikleri sırasıyla sabit örneklem genişliği testlerinde 178, Pocock testinde 210 ve O'Brien & Fleming testinde ise 182 olarak hesaplanmaktadır. Bu çalışmanın popülasyonda çok az

rastlanan bir hastalıkla ilgili olduğu ve her yıl ancak 50 kişide bu hastalığın gözlemlendiği varsayılırsa maksimum çalışma süresi, sabit örneklem genişliği testi için 3,56 yıl olarak bulunurken, O'Brien & Fleming testi için 3,64 yıl ve Pocock testi için 4,20 olarak bulunacaktır. Örneklem genişliğindeki azalma şansına rağmen, bir araştırmacı, sabit örneklem genişliği deneme düzeni ile yaklaşık bir yıl daha önce bir sonuç elde edebileceğini garanti ettiğinden, 4,20 yıl bekleme olasılığı nedeniyle Pocock testini kullanmak istemeyebilir.

Kullanılacak test tipini seçmek kadar, gözlemlerin alınacağı grupların sayısına da karar verilmelidir. Bazı durumlarda 1 veya 2 ara analiz yapılması bile örneklem genişliğinin azalmasında etkili olabilmektedir ve genellikle 4 veya 5 ara analizle kurulan deneme düzenleri genellikle yeterli olmaktadır. Etki büyüklüğünün yüksek değerlerinde O'Brien & Fleming testinin karar verene kadar kullanacağı örneklem genişliği, yapılması planlanan toplam analiz sayısı (K) arttıkça azalmaya devam ederken, Pocock testi için tekrar artmaya başlamaktadır. Aynı zamanda etki büyüklüğü sıfıra yakın olduğunda, Pocock testi için de maksimum örneklem genişliği ve beklenen örneklem genişlikleri analiz grubu sayısı ile birlikte anlamlı şekilde artmaktadır. Buna göre, O'Brien Fleming testinin kullanılacağı bir grup ardışık deneme düzeni için $K = 10$ kadar analiz grubu seçmek optimum durumun meydana gelmesini sağlarken, Pocock testinin kullanılacağı bir grup ardışık deneme düzeninde $K = 5$ analiz grubundan daha fazla kullanmak çok da mantıklı görünmemektedir.

Sonuç olarak bu iki grup ardışık test tipinin kendi avantaj ve dezavantajları vardır. Etki büyüklüğü yüksek olduğunda Pocock testi, O'Brien & Fleming testinden daha az örneklem genişliği kullanarak erken sonlandırma avantajına sahiptir. Buna karşın etki büyüklüğü az olduğunda Pocock testi oldukça büyük maksimum örneklem genişliğine sahiptir ve bu değer sabit örneklem genişliği testinde gerek duyulandan çok daha büyüktür. İlk ara analizlerde, Pocock testinin tespit edemeyeceği kadar küçük etki büyüklüklerini O'Brien & Fleming testi de tespit edemeyecektir. Ancak daha sonraki analizlerde O'Brien & Fleming testi azalan kritik değerleri sayesinde bu etki büyüklüğünü tespit edebilecek duruma gelirken Pocock testi belki de son analize kadar devam edecek ve denemeyi H_0 hipotezini kabul ederek sonlandıracaktır. Aynı zamanda

O'Brien & Fleming testi, düşük maksimum örneklem genişliği avantajının yanında, eğer son analize kadar deneme devam ederse sabit örneklem genişliği deneme düzeni testine oldukça yakın olan kritik değerlere sahip olmak avantajına da sahiptir.

Bazı çalışmalarda, araştırmacılar yeni muamelenin birincil etkisinin yanında muamele hakkında daha fazla bilgi edinebilmek için örnekleme devam etmek isteyebilirler. Yeni bir muamelenin etkinliğinin sadece bir çalışmada kanıtlanması, etkinliğinin gösterilmesinden hemen sonra gerekli şekilde onaylanması için yeterli değildir. Bunun yanında, muamelenin standart bir muamele veya plaseboya göre güvenliğini göstermek için daha fazla veri gerekmektedir. Ayrıca başka denemeler yapılarak diğer deneysel koşullardaki etkinlik durumunun incelenip aynı etkinlik kanıtının tekrarlanıp tekrarlanmadığını incelemek gerekli olabilir. Bu durum özellikle, birçok kanser tipi gibi, etkin tedavilerin mevcut olduğu ve yeni muameleler önermenin de olası olduğu, çok küçük ilerlemeleri tespit etmek için büyük örneklem genişliklerinin gerektiği hastalıklar için olan muamelelerde yaygındır. Böyle durumlarda, araştırmacıların amacı, gerçekten bir muamele üzerine mümkün olduğu kadar fazla bilgi toplamaktır ve gerektiğinde denemeyi erken durdurarak hastaları açık şekilde etkin olmayan bir muameleye atamayı bırakmak için en önemli neden etik gerekliliktir.

Grup ardışık deneme düzenlerinin erken aşamalarında yüksek bir test istatistiği (Z) gözleendiğinde, bu değer durma sınırı ile kesişmemiş olsa bile VGİK etik nedenlerle bir denemeyi durdurmak isteyebilir. Örneğin; durma sınırıyla kesişmese bile $Z = 4$ gibi büyük bir test istatistiği hesaplandığında VGİK denemenin durdurulmasını zorunlu hale getirebilir. Bu nedenle herhangi bir çalışmanın ilk ara analizinden önce yapılan ilk toplantısında VGİK, testin dikkate aldığı kritik değer dışında denemeyi erken sonlandırmak için dikkate alınacak kriterlerin ne olduğunu tartışmalı ve bu kriterlere karar vererek denemenin başında belirlemelidir. Bununla birlikte, elbette ki tüm olası durumlar öngörülemeyecek ve beklenmeyen sonuçlar ortaya çıkacaktır. Bu gibi durumlarda da VGİK, bu sonucun gözleendiği ara analizden sonraki ilk toplantısında bu durum hakkında karar vermelidir.

Sonuç olarak bir grup ardışık deneme planlanırken kullanılacak test tipine karar vermek için çalışılan konunun ne olduğu, kolaylıkla denek bulunup bulunamayacağı ve erken sonlandırma ya da daha küçük muamele farklılıklarını test etme durumlarından hangisiyle ilgilenildiği önem kazanmakta ve kullanılacak test tipi bu kriterlere göre farklılık göstermektedir. Denek bulmanın çok zor olduğu ya da yüksek risk nedeniyle daha küçük bir örneklem genişliğiyle çalışılması gereken durumlarda, ilk ara analizlerde daha küçük muamele farklılıklarını tespit etmeyi sağlayan kritik değerlere sahip olan testler tercih edilmelidir.

KAYNAKLAR

1. **DeMets D.L.**, Sequential Designs in Clinical Trials, *Cardiac Electrophysiology Review*, 1998(2):57-60.
2. **Chow S.C, Liu J.P.**, Design and Analysis of Clinical Trials: Concepts and Methodologies, *Wiley Series in Probability and Statistics*, Second Edition, United States of America, 2004.
3. **Jennison C., Turnbull B.W.**, Group Sequential Methods with Applications to Clinical Trials, *Chapman & Hall/CRC*, United States of America, 2000.
4. **O'Brien P. C.**, Data and Safety Monitoring, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(2):1362-1371.
5. **Yiannoutsos C.T.**, Clinical Trials: A Short Course, Eriřim: <http://www.biostat.iupui.edu/ShortCourse/clinical2.pdf>, Eriřim Tarihi: 29.07.2005.
6. **Ederer F.**, Clinical Trials, History of, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(2):864-874.
7. **Heilbrun L. K.**, Clinical Trials: Part 1 of 2, Eriřim: <http://www.med.wayne.edu/pdfs/Clinical%20Trials%201.pdf>, Eriřim Tarihi: 27.06.2006.
8. **Heilbrun L. K.**, Clinical Trials: Part 2 of 2, Eriřim: <http://www.med.wayne.edu/pdfs/Clinical%20Trials%202.pdf>, Eriřim Tarihi: 27.06.2006.
9. **Lesaffre E., Verbeke G.**, Clinical Trials and Intervention Studies, Eriřim: <http://www.wiley.com/legacy/wileychi/eosbs/pdfs/bsa099.pdf>, Eriřim Tarihi: 27.06.2006.
10. **Todd S., Stallard N.**, A New Clinical Trial Design Combining Phases 2 and 3: Sequential Designs with Treatment Selection and a Change of Endpoint, *Drug Information Journal*, 2005(39):109-118.
11. **Meinert C. L.**, Clinical Trials, Overview, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(2):875-890.
12. **Shapiro S., Rosenberg L.**, Bias in Case-Control Studies, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(1):390-399.
13. **Hill H. A., Kleinbaum D. G.**, Bias in Observational Studies, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(1):410-416.

14. **Bryant J., Cronin W., Wieand S.**, Multicenter Trials, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(5):3354-3360.
15. **Gail M. H.**, Controls, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(2):1159-1160.
16. **Taşdelen B., Kanık E.A.**, Klasik Hipotez Kontrollerine Ait Örnek Genişliklerinin Belirlenebilmesi için Gerekli Formüller ve Tablolar, *Mersin Üniversitesi Tıp Fakültesi Dergisi*, 2004(4):438-446.
17. **Lachin J. M.**, Sample Size Determination, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(7):4693-4704.
18. **Lakatos E.**, Sample Size Determination for Clinical Trials, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(7):4704-4711.
19. **Aplenc R., Zhao H., Rebbeck T.R., ProPERT K.J.**, Group Sequential Methods and Sample Size Savings in Biomarker-Disease Association Studies, *Genetics*, 2003(163):1215-1219.
20. **Van der Tweel I.**, Sequential analysis in clinical trials and epidemiological studies, Erişim Adresi: www.bio.uu.nl/~biostat/seminar%20JC.ppt, Erişim Tarihi: 29.07.2005.
21. **Ellenberg S. S.**, The What, Why and How of Clinical Trial Data Monitoring Committees, Erişim Adresi: http://www.uab.edu/ethicscenter/2_26_00uab.ppt, Erişim Tarihi: 27.06.2006.
22. **Kanık E.A., Sungur M.A.**, Klinik Denemelerde Hata Kaynakları ve Yanlılık, *Mersin Üniversitesi Tıp Fakültesi Dergisi*, 2005(3):287-292.
23. **Dorak M.T.**, Bias and Confounding, Erişim Adresi: <http://www.dorak.info/epi/bc.html>, Erişim Tarihi: 06.07.2005.
24. **Delgado-Rodriguez M. and Llorca J.**, Bias, *J. Epidemiol. Community Health*, 2004(58):635-641.
25. **Dorak M.T.**, Error, Bias and Confounding, Erişim Adresi: <http://www.emro.who.int/tdr/ProposalDevelopment2002/Presentations/5a.%20Error.%20Bias.%20Confounding.ppt>, Erişim Tarihi: 05.07.2005.
26. **Greenland S.**, Confounding, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(2):1118-1125.

27. Randomisation, Erişim:
<http://web1.iop.kcl.ac.uk/iop/Departments/PsychMed/CTU/Randomisation.shtml>, Erişim Tarihi:
20.08.2003.
28. **Wittes J.**, Randomized Treatment Assignment, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(6):4486-4494.
29. **Day S. J.**, Blinding or Masking, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(1):518-525.
30. **Whitehead J.**, Sequential Methods for Clinical Trials, Armitage P., Colton T. (Ed.), *Encyclopedia of Biostatistics*, Wiley, Second Edition, United Kingdom, 2005(7):4896-4905.
31. **Mazumdar M., Liu A.**, Group Sequential Design for Comparative Diagnostic Accuracy Studies, *Statistics in Medicine*, 2003(22):727-739.
32. **Mazumdar M.**, Group Sequential Design for Comparative Diagnostic Accuracy Studies: Implications and Guidelines for Practitioners, *Medical Decision Making*, 2004(24):525-533.
33. **Chen K.**, Optimal Sequential Designs of Case-Control Studies, *The Annals of Statistics*, 2000(5):1452-1471.
34. **Kanık A.**, Örneklem Dağılımı ve Hipotez Kontrolleri, *Basılmamış Ders Notları*.
35. **Chow S.C., Shao J., Wang H.**, Sample Size Calculations in Clinical Research, *Marcel Dekker, Inc.*, United States of America, 2003.
36. **Müller H.H., Schafer H.**, Adaptive Group Sequential Designs for Clinical Trials: Combining the Advantages of Adaptive and of Classical Group Sequential Approaches, *Biometrics*, 2001(57):886-891.
37. Cytel's East Clinical Trial Design System, Erişim:
http://www.cytel.com/Products/East/Example_images, Erişim Tarihi: 29.07.2005.
38. Data and Safety Monitoring Report: Closed Session Efficacy Comparisons (Level 3), Erişim:
http://icsc.org/ReportLevel3_Jul2003.doc, Erişim Tarihi: 04.08.2005.
39. Microsoft Office Standard Edition for Students and Teachers, Microsoft ® Office Excel 2003 (11.8134.8132) SP2, 1983-2003 Microsoft Corporation, <http://office.microsoft.com/tr-tr/default.aspx>.
40. **Karrison T.G., Huo D., Chappell R.**, A Group Sequential, Response-Adaptive Design for Randomized Clinical Trials, *Controlled Clinical Trials*, 2003(24):506-522.

ÖZGEÇMİŞ

30.03.1980 tarihinde Kayseri’de doğdu. İlk, orta ve lise öğrenimlerini Kayseri’de tamamladıktan sonra 1999-2000 Eğitim-Öğretim yılında lisans öğrenimine başladığı Mersin Üniversitesi Fen-Edebiyat Fakültesi Biyoloji Bölümü’nden 20.06.2003 tarihinde mezun oldu. 2003-2004 Eğitim-Öğretim yılında Mersin Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı’nda bir yıl süreyle özel öğrenci olarak yüksek lisans öğrenimi gördükten sonra, 2004-2005 Eğitim-Öğretim yılında Mersin Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı’nda yüksek lisans öğrenimine başladı. Halen Mersin Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı’nda Araştırma Görevlisi olarak çalışmaktadır.