

İki Evreli Faz II Klinik Denemesi ve Farklı Tasarımlara Göre Örneklem Büyüklüğü

Two-Stage Phase II Trials and Sample Size for Different Designs: Review

Güven ÖZKAYA,^a
Bülent EDİZ,^a
İlker ERCAN^a

^aBiyostatistik AD,
Uludağ Üniversitesi Tıp Fakültesi,
Bursa

Geliş Tarihi/Received: 02.11.2011
Kabul Tarihi/Accepted: 28.11.2011

Yazışma Adresi/Correspondence:
Güven ÖZKAYA
Uludağ Üniversitesi Tıp Fakültesi,
Biyostatistik AD, Bursa,
TÜRKİYE/TURKEY
guvenozkaya@gmail.com

ÖZET Günümüzde insanlığı tehdit eden birçok hastalık mevcut iken yenileri de eklenmektedir. Hastalıklara karşı uygulanacak tedaviyi aramak veya hastalıkların ilerlemelerini durdurmak tıp biliminin en büyük amaçlarındandır. Yeni ilaç geliştirme maliyetli ve uzun bir süreçtir ve tıbbi bir gereksinim olarak hastalığa karşı etkili bir ilaç geliştirme uzun bir zaman gerektirir. Ciddi protokoller geliştirmek, iyi bir araştırma ekibi kurmak ve önemli bir bütçeye sahip olmak gerekmektedir. Faz II çalışmaları, ilacın kontrollü klinik çalışmalarının ilkidir ve birkaç yüz hastadan fazlasını içermezler. Faz II çalışmalarının amacı, çalışmada yer alan hastalardaki özel belirti veya belirtiler için klinik bitiş noktasına dayanan ilacın etkinliğini ölçmekle birlikte doz aralığına, faz III çalışmaları için dozlara, ortak kısa dönem yan etkilere ve ilaçla ilişkili olan risklere karar vermektir. Etik ve ekonomik nedenlerle tasarlanan çok evreli faz II klinik denemelerden en popüler olanı iki-evreli denemelerdir. İki-evreli faz II klinik denemelerden Simon'un geliştirdiği optimal ve minimaks tasarımları, Kepner/Chang'ın tasarımı ve SWOG tasarımı incelenmiştir. Faz II klinik deneme tasarımlarında $H_0:p \leq p_0$ sıfır hipotezine karşı $H_1:p \geq p_1$ alternatif hipotezi test edilmektedir. Faz II klinik deneme tasarımlarından optimal, minimaks ve Kepner/Chang tasarımları α ve β hata oranları bakımından birbirlerinden çok büyük farklılık göstermemektedir. SWOG tasarımı ise bu iki hata oranı diğer tasarımlara kıyasla farklılıklar göstermektedir.

Anahtar Kelimeler: Klinik denemeler; iki evreli faz II denemeler; örneklem büyüklüğü; I. tip hata; II. tip hata

ABSTRACT While many diseases have threatened humanity nowadays, the newer diseases are added to the list. The ultimate objectives of medicine science are researching treatment against to disease or stopping progress of disease. New drug development is cost and has long process, and requires a long time as a medical need for effective drug development against to the disease. Developing serious protocols need establishing a research group and having a significant budget. Phase II studies are the first controlled clinical studies of the drug, and they involve no more than several hundred patients. The primary objectives of phase II studies are not only to initially evaluate the effectiveness of a drug based on clinical endpoints for a particular indication or indications in patients with the disease or condition under study but also to determine the dosing ranges and doses for phase III studies and the common short-term side effects and risks associated with the drug. The most popular one of the multi-stage phase II clinical trials, designed considering ethical and economic reasons, has two stages. Simon's optimal design, minimax design, Kepner/Chang's design and SWOG design are investigated as two-stage phase II clinical trials. $H_0:p \leq p_0$ null hypothesis is tested against $H_1:p \geq p_1$ alternative hypothesis in phase II clinical trials. Phase II clinical trials of Simon's optimal design, minimax design and Kepner/Chang designs are not too dissimilar according to α and β error rate. However, SWOG design is more different than others for both error rates.

Key Words: Clinical trials; two-stage phase II trials; sample size; type I error rate; type II error rate

Günümüzde insanlığı tehdit eden birçok hastalık mevcut iken yenileri de eklenmektedir. Hastalıklara karşı uygulanacak tedaviyi aramak veya hastalıkların ilerlemelerini durdurmak tıp biliminin en büyük amaçlarından biridir. Bu amaçla ilaç geliştirme ve geliştirilen ilaçların insanlarda kullanımını öncesi yapılan çalışmalar belirli koşullar altında yürütülmektedir. Bu araştırmalarda uyulması gerekli kurallar konusunda uluslararası anlayış birliği geliştirilmiştir.¹ Çalışılan ürünler ve tedavi protokolleri tıp bilimi alt dallarının ve tıp dışındaki diğer bilim dallarının işbirliğiyle klinik denemeler vasıtasıyla incelenir. Bu çalışmalarda insan üzerinde yapılan biyomedikal araştırmaların hasta ve denek hakları açısından etik ilkelere uygun olup olmadığı etik kurullar tarafından titizlikle incelenir ve denetlenir.²

Klinik denemeler, sağlık çalışanlarının klinikte gerçekleştirdiği klinik araştırmalardır.³ Farklı zamanlarda farklı bireyler ve organizasyonlar tarafından klinik deneme için farklı tanımlamalar yapılmıştır. Meinert, aynı zaman dilimi içerisinde kaydedilen tedavi gören ve izlenen hastaların yer aldığı; bir grupta test tedavisi ile tedavi gören hastaların, diğer bir grupta kontrol tedavisi alan hastaların sonuçları karşılaştırılarak insanlardaki tedavinin etkinliğinin değerlendirilmesi için tasarlanmış, planlanmış bir deney olarak tanımlamıştır.⁴ Bu tanım, bir klinik denemenin tedavinin etkinliğini ölçmesinde kullanıldığını gösterir. Piantadosi, klinik denemeleri tıbbi tedavinin insanlar üzerindeki deneysel testi olarak tanımlamıştır.⁵ Spilker, klinik denemeleri, faz I, faz II ve faz III ilaç araştırmalarının ölçümünü klinik çalışmaların bir alt kümesi olarak düşünmüştür.⁶ Klinik çalışmalar, tıbbi hastalıkların engellenmesi, tanı teknikleri ve tedavilerin değerlendirilmesi için bütün bilimsel yaklaşımları içeren sınıfta bulunur.

KLİNİK GELİŞMELERİN FAZLARI

Yeni ilaç geliştirme maliyetli ve uzun bir süreçtir ve tıbbi bir gereksinim olarak hastalığa karşı etkili bir ilaç geliştirme uzun bir zaman gerektirir. Ciddi protokoller geliştirmek, iyi bir araştırma ekibi kurmak ve önemli bir bütçeye sahip olmak gerekmektedir. İlaç geliştirme sürecinde ilk olarak molekül

çalışmaların yapılmaktadır. Yeni molekül geliştirme, gerekli analiz ve sentezlerin yapılması gibi aşamalardan sonra, klinik öncesi fazda potansiyel yeni ilacın canlı bir organizmada öldürücü dozu, gebeliğe etkileri gibi toksikoloji testleri gerçekleştirilmektedir. Bu aşamada etkinliği kanıtlanan ilacın klinik denemelerine geçilmektedir. Dört fazda gerçekleştirilen klinik denemelerde insanlar kullanılmaktadır.

Faz I klinik araştırması, insanlarda kullanılacak yeni ilaç araştırmalarının başlangıcı niteliğinde bir araştırma sağlar.³ Faz I denemeleri, tedavinin maksimum tesirli dozu hakkında bilgi sağlar. Faz I klinik araştırmalarının birincil amacı iki parçadır. Birincisi, ilacın insanlar üstündeki metabolizmasını ve farmakolojik aktivitelerini, artan dozla ilişkili olarak yan etkileri ve etkinlik üzerine erken kanıtları belirlemektedir.³ Diğer amacı; iyi kontrol edilmiş ve bilimsel olarak faz II çalışmalarına uygun tasarımlara izin veren ilacın farmokinetik ve farmakolojik etkileri hakkında yeterli bilgi edinmektir. Böylece faz I klinik araştırmaları, ilaç metabolizması, biyolojik uygunluk, doz aralığı, çoklu doz çalışmalarını içerir.³ Faz I denemeleri genellikle her bir doz seviyesi için 3' ten 6'ya kadar denek üzerinde çalışır.⁷ Faz I klinik araştırmaları 20 ile 80 hasta veya normal gönüllü içerir.³ Faz I çalışmaları için genel protokol daha sonraki faz çalışmalarından daha az detaylı ve daha esneklerdir. Çünkü bu çalışmalar araştırmanın taslağını oluşturur.

Faz II çalışmaları, ilacın kontrollü klinik çalışmalarının ilkidir ve birkaç yüz hastadan fazlasını içermezler. Faz II çalışmalarının amacı, çalışmada yer alan hastalardaki özel belirti veya belirtiler için klinik bitiş noktasına dayanan ilacın etkinliğini ölçmekle birlikte doz aralığına, faz III çalışmaları için dozlara, ortak kısa dönem yan etkilere ve ilaçla ilişkili olan risklere karar vermektir. Klinik araştırmalar birkaç yüz hastadan fazlasını içermemesine rağmen, genişletilmiş faz II klinik çalışmaları birkaç bin hasta içerebilir. Farmasötik şirketler, bu fazları faz IIA ve IIB olarak ayırmışlardır. Dozu ölçmek için dizayn edilen klinik çalışmalar faz IIA olarak nitelendirilirken, ilacın etkinliğine karar vermek için dizayn edilen çalışmalar faz IIB olarak nitelendirilir.³

Faz III çalışmaları genişletilmiş kontrollü ve kontrollü olmayan denemelerdir. Faz III çalışmaların ana amacı, ilacın tüm yarar ve risk ilişkilerini ölçmek için gerekli etkinlik ve güvenlik hakkında ek bilgi toplamakla birlikte ilaçların sınıflandırılması için hekimlere uygun temeli sağlamaktır. Birkaç yüz ile birkaç bin arası hasta içerebilen faz III çalışmaları, ilacın etkinliği hakkındaki ilk kanıt elde edildikten sonra gerçekleştirilir.³

İlaç geliştirmelerinde faz I çalışmaları klinik farmakolojinin erken evresi olarak nitelendirilirken faz II ve faz III çalışmalar ise klinik gelişmenin sonraki evreleri olarak nitelendirilir.

Faz IV denemeleri bir ilacın pazarlanması için uygun bulunmasından sonra gerçekleşen çalışmalar olarak nitelendirilir. Faz IV çalışmalarındaki amaç, sonraki ters reaksiyonların sonucunu açıklamak ve ilacın morbidite ve mortalite üzerindeki etkisine karar vermektir. Faz IV denemesi hasta popülasyonu üzerinde yapılan çalışmalardır. Faz IV çalışmaları, rakip ürünlerle karşılaştırılma yapılmasına olanak sağlayan çalışmalardır.

FAZ II KLİNİK DENEMELERİ

Faz II klinik denemelerin ana amacı, Chang ve Herson'a göre yeni tedavilerin terapötik etkinliğini ve toksisitesini belirlemektir.^{8,9} Herson ve Bryant, faz II denemelerinin en büyük işlemsel amacının "n" hastadaki cevap oranını tahmin etmek olduğunu belirtmiştir.^{9,10} Simon'a göre "yeni bir kanser önleyici ilaçta faz II denemesinin amacı ilacın belirli bir tipteki tümöre karşı yeterli gelişmeyi sağlayıp sağlamadığını belirlemektir".⁷

Kanser tedavisinin faz II çalışması, anti-tümör tedavi etki derecesinin başlangıç tahminini elde etmek için yapılan kontrolsüz bir denemedir.⁷ Maksimum dozda, maksimum etki için birçok kanser tedavisinin bırakılması gerektiğinden, faz I denemeleri tedavinin maksimum tesirli dozu hakkında bilgi sağlar. Sonuç olarak, bu tip denemeler anti-tümör aktivitesi hakkında ya çok az ya da hiç bilgi sağlamamaktadır. Tümörlerinde en azından %50 azalma gözükten hastaların oranı birçok faz II çalışmasının öncül bitiş noktasıdır.⁷ Bunlar kontrolsüz denemelerdir ve hastalığın

tedavisindeki ilacın rolünü veya tedavinin etkinliğine bu denemelerle karar verilmez. Yeni bir anti-kanser ilacının faz II denemesinin amacı, ilacın belirli tipteki bir tümöre karşı daha fazla bir gelişmeyi garanti etmek için yeterli aktiviteye sahip olup olmadığını kararlaştırmaktır. Daha fazla gelişme, ilacın diğer ilaçlarla birleştirilmesini, hastalığı daha az ilerlemiş olanların değerlendirilmesi veya standart tedavi için yaşam sonuçlarının kararlaştırıldığı faz III çalışmasının başlaması anlamına gelebilir. İlacın ana maddelerinin kombinasyonunun faz II denemesi, tedavinin standart terapiye karşı temel kontrollü klinik değerlendirmeyi garanti etmesini yeterince taahhüt edebilme durumuna göre karar vermek için yürütülür.⁷

Faz II klinik deneme tasarımları tek-evreli ve çok evreli testlerle incelenebilir. Shuster ve Green, ikiden fazla evre ile çalışılmasının pratik olmayacağı belirtmiştir.^{11,12} Jung ve ark., etik ve ekonomik nedenlerle dizayn edilen çok evreli klinik denemelerden en popüler olanının iki-evreli olduğunu belirtmiştir.¹³ Simon ikiden fazla evreli faz II klinik denemelerin çok merkezli olarak yürütülmesinin zor olacağını belirtmiştir.⁷ Chang ise faz II klinik denemelerinde maksimum evre sayısını iki veya üç olarak uygun olacağını belirtmiş ve birçok pratik faz II testlerinde sıkça test yapılmasının gerekmediğini vurgulamıştır.⁸ Etkinliğin testi için daha fazla evreli test gerekmesine rağmen iki-evreli tasarım dizisel tasarımlarda yeterli başarıyı sağladığı düşünülmektedir.¹⁴⁻¹⁶ Bu nedenlerden dolayı literatürlerde iki-evreli faz II çalışmaları çoğunluktadır.

Faz II klinik deneme tasarımlarında $H_0: p \leq p_0$ sıfır hipotezine karşı $H_1: p > p_1$ alternatif hipotezi test edilmektedir. p_0 , geliştirilen ilacın reddedileceği en yüksek cevap oranıdır. Bir başka deyişle, kötü tedaviye verilen en yüksek cevap oranıdır. p_1 , geliştirilen ilacın kabul edilebilmesi için en düşük cevap oranıdır. Bir başka deyişle, iyi tedaviye verilen en düşük cevap oranıdır. H_0 hipotezinin kabul edilmesi durumunda geliştirilen ilaç bir sonraki klinik denemeye geçemez. H_1 hipotezinin kabul edilmesi durumunda bir sonraki klinik denemeye geçilir. $\alpha = P(H_0 \text{ ret} | H_0 \text{ doğru})$ ve $\beta = P(H_1 \text{ ret} | H_1 \text{ doğru})$ hata oranları kısıtları altında örneklem büyüklüğünün minimum yapılması istenir.

Etik nedenlerden dolayı ajan için yapılan denemenin etkisiz olduğunu göstermek için erken sonlandırılabilir.^{12,17} Herson, istatistiksel nedenlerle erken sonlandırma kuralının yararlı olduğunu belirtmiştir.⁹ Erken sonlandırma kuralı faz II denemelerinde etkisiz ilaca veya terapiye maruz kalan hasta sayısını azaltmak amacıyla çalışmanın sonlandırılması şeklinde uygulanan bir kuraldır.⁹

İKİ-EVRELİ FAZ II KLİNİK DENEME DİZAYNLARI

Birçok tasarım geliştirilmekle birlikte başlıcaları aşağıda belirtilmiştir:

- Simon'ın Optimal İki-Evrelî Faz II Klinik Deneme Tasarımı
- Minimaks İki-Evrelî Faz II Klinik Deneme Tasarımı
- Kepner/Chang İki-Evrelî Faz II Klinik Deneme Tasarımı
- Southwest Onkoloji Grubu Standart Tasarımı

SİMON OPTİMAL VE MİNİMAKS İKİ-EVRELİ FAZ II KLİNİK DENEME TASARIMI

Simon'un tasarımında optimal olduğu düşünülen iki-evrelî tasarımlar sunulur. α ve β hata kısıtları altında eğer ilacın ana maddesi düşük aktiviteye sahip ise beklenen örneklem büyüklüğü minimum olur.

Çalışmanın birinci ve ikinci evresindeki hasta sayıları n_1 ve n_2 şeklinde gösterilirken beklenen örneklem büyüklüğü (expected sample size) $EN=n_1+(1-PET)n_2$ (PET, 1. evreden sonraki "erken sonlandırma olasılığını" (PET, probability of early termination) göstermektedir). İlk evreden sonra çalışmanın sona erip ermeyeceğinin kararı birinci evrede n_1 hastada gözlenen cevap sayısına bağlıdır. Beklenen örneklem büyüklüğü EN ve erken sonlandırma olasılığı, "p" gerçek cevap olasılığına bağlıdır. Eğer r_1 sayıda veya daha az cevap gözlenmesi durumunda deneme ilk evrede sona erecektir. Birinci evre sonundaki cevap sayısı ölçütüdür r_1 'dir. Bu olay $PET=B(r_1; p, n_1)$ olasılığı ile gerçekleşir (B, kümülatif binomial dağılımı gösterir). Eğer r sayıda veya daha az cevap gözlenirse, ilaç ikinci evre so-

nunda reddedilecektir. r, ikinci evre sonundaki toplam cevap sayısı ölçütüdür. Böylece p başarı olasılığı ile birlikte bir ilacı reddetme olasılığı

$$B(r_1; p, n_1) + \sum_{x=r_1+1}^{\min(n_1, r)} b(x; p, n_1)B(r-x; p, n_2) \quad (1)$$

şeklinde gösterilir.⁵

p: Gerçek cevap olasılığı

x: raslantı değişkeni

r_1 : 1. evre sonundaki cevap sayısı ölçütü

r: 2. evre sonundaki toplam cevap sayısı ölçütü

n: Denemedeki toplam denek sayısı

n_1 : 1.evredeki denek sayısı

n_2 : 2.evredeki denek sayısı

b: Binomial olasılık yoğunluk fonksiyonu

Bu tasarım yaklaşımı -cevap olasılığı p_0 iken p_0, p_1, α ve β belirlenmesi ve sonra hata olasılığı kısıtlarını sağlayan ve beklenen örneklem büyüklüğünü minimize eden iki-evrelî tasarıma karar verilmesini dikkate alır.⁷ Optimizasyon, r_1 ve r değerleri gibi n_1 ve n_2 'nin tüm değerleri alınarak yapılır. Simon'ın tasarımında ilacın erken kabul edilmesine izin verilmemektir.⁷ İlaç düşük aktiviteye sahip olduğunda erken sonlandırma için etik zorunluluk oluşmaktadır. İlaç tatmin edici aktiviteye sahip olduğunda ($\geq p_1$), cevapların oranını, boyutunu ve dayanıklılığını tahmin etmek için ilave hastalıklarla sıklıkla çalışılır.⁷ Cevap olasılığı p_0 iken, tasarımın beklenen örneklem büyüklüğünü en küçük tutarak optimize edilmektedir. Bu durum etik nedenlerden kaynaklanmaktadır. Eğer ilk evredeki cevap sayısı, son karar kriterini (r) aşarsa, erken sonlandırma ile beklenen örneklem büyüklüğü azaltılabilir. Bu durum H_0 hipotezi altında, performans üzerinde önemsiz bir etkiye sahiptir. Eğer ilgilenilen, H_0 'ın erken reddi ise her nasılsa, daha az katı erken reddetme kuralı kullanılmaktadır.

Simon'ın minimaks tasarımı α ve β hata kısıtları altında beklenen örneklem büyüklüğünü minimum yapan tasarımdır. p_0, p_1, α ve β 'nin belirlenmiş değerleri için kesin binomial olasılıklar kullanılarak optimal tasarıma karar verilir. Toplam örneklem büyüklüğünün her bir değeri

için (n) ve $(1, n-1)$ aralığındaki n_1 'in her bir değeri için $p=p_0$ iken iki kısıtı sağlayan ve beklenen örneklem büyüklüğünü minimize eden r_1 ve r 'nin tam sayı değerlerine karar verilir. $r_1 \in (0, n_1)$ aralığında araştırılarak bulunmuştur. r_1 'in her bir değeri için β hata kısıtını sağlayan maksimum r değerine karar verilir. Daha sonra (n, n_1, r_1, r) parametre kümesinin α hata kısıtını sağlayıp sağlamadığı incelenir. Eğer sağlıyorsa, beklenen örneklem büyüklüğü bir önceki uygulanabilir tasarım tarafından başarılan minimumla karşılaştırılır ve r_1 üzerinden araştırma devam eder. n 'i sabit tutarak, maksimum örneklem büyüklüğü n için optimal iki-evreli tasarımı bulmak amacıyla n_1 'in dağılım aralığı incelenir.

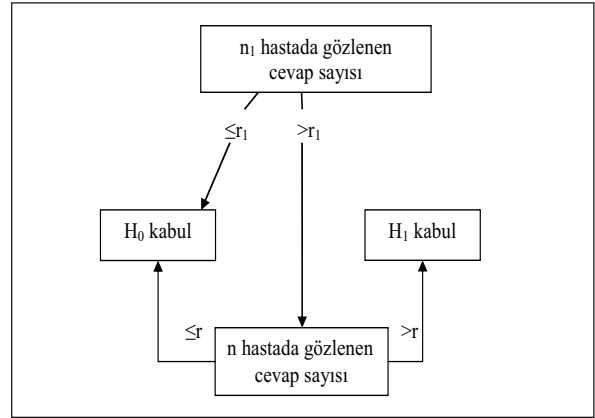
n üzerindeki araştırmada aralığın düşük değeri aşağıdaki gibi oluşur.

$$\bar{p}(1-\bar{p}) \left[\frac{z_{1-\alpha} + z_{1-\beta}}{p_1 - p_0} \right]^2$$

Bu noktadan başlayarak, hata olasılıkları kısıtlarını sağlayan ve $n_1, n_2 > 0$ olduğu iki-evreli tasarımda belirlenen maksimum örneklem büyüklüğünün (n) en küçük olduğunu garanti etmek için kontrol edilmiştir. Sayım prosedürünün incelenmesi n 'in minimum değerinden yukarı doğru, optimum değere karar verilene kadar devam eder. Binomial dağılımın kesikli olmasından, sabit bir n için minimum beklenen örneklem büyüklüğü n 'in unimodal fonksiyonu değildir. " n " artarken yerel minimum değeri artmaktadır ve burada global minimum değerinin bulunduğu açıktır. Simon'un her iki tasarımının çalışma şekli aşağıda verilmiştir (Şekil 1).

KEPNER/CHANG İKİ-EVRELİ FAZ II KLİNİK DENEME TASARIMI

Kepner/Chang iki-evreli tasarımında, terapiye objektif cevap verenlerin oranı c ile gösterilmektedir.¹⁸ c , binomial olasılık dağılımına sahiptir ve büyük değerlere ulaşmasıyla H_0 reddedilir. α ve β kısıtlarını sağlayan bir ret bölgesi olan ve en küçük örneklem büyüklüğü olan n , kesin-tek-evreli örneklem büyüklüğü olarak adlandırılır. k -evreli tasarımlar tipine göre sınıflandırılır.¹⁸ Tip 1 tasarımlar, etkinlik hakkında sonuca varmak için



ŞEKİL 1: Simon'un iki-evreli faz II klinik deneme optimal ve minimaks tasarımlarının adımları.

r_1 : ilacın reddedilmesi ve denemenin sonlandırılması için gerekli olan ilaca cevap veren maksimum hasta sayısı

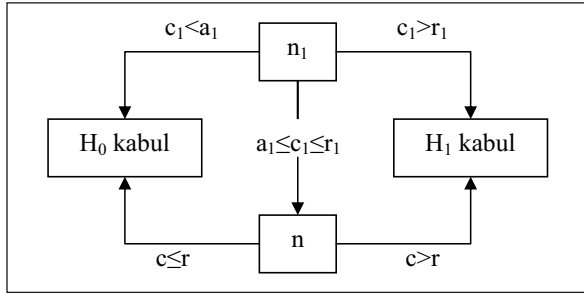
n_1 : 1. evrede incelenen hasta sayısı

r : Deneme sonunda ilacın reddedilmesi ve denemenin sonlandırılması için gerekli olan ilaca cevap veren maksimum hasta sayısı

n : Deneme boyunca incelenen toplam hasta sayısı

erken sona erer. Tip 2 tasarımlar, yararlılık hakkında sonuca varmak için erken sona ererken, Tip 3 tasarımlar yararlılık ya da etkinlik hakkında sonuca varmak için erken sona erer. Kepner ve Chang'a ait temel teoremler, n örneklem büyüklüğüne sahip tek-evreli testin anlamlılık seviyesi ve gücü ile aynı olan, her tipe ait k -evreli kesin grup ardaşık tasarımların varlığını garanti eder $(1, 2, \dots, k)$ -evrelerinde sırasıyla n_1, n_2, \dots, n_k örneklem büyüklüğü.¹⁸ Örneklem büyüklükleri için kısıtlar $n_1 \geq n_2 \geq \dots \geq n_k$ ve $N = \sum n_i = n$ şeklindedir. Ayrıntılı bir araştırma yapılarak, her tipin $N \leq n$ durumunu gerçekleştiren α ve β koşulunu sağlayan bir tasarım bulunabilir.

Pratikte, 2 ya da 3 aşamalı tasarımlar, faz II ve pilot çalışmalar için genellikle yeterlidir. Her üç tasarım tipi için $k \leq 3$ ve kullanıcıdan alınan giriş parametreleri $p_0, \Delta = (p_1 - p_0) > 0, a, 1 - \beta$ için Kepner/Chang detaylı bir araştırma ile en küçük N örneklem büyüklüğüne sahip tasarımı bulmayı hedeflerken; sadece bahsedilen kısıtlamaları sağlamanın yanında ayrıca da n_1, n_2, \dots, n_k örneklem büyüklüklerinin en aza indirilmesini gerektiren bir yazılım yazmışlardır. $k=2$ olduğu durumda tip 1 tasarımın nasıl bulunduğunu şekilde verilmiştir (Şekil 2). Kepner/Chang tasarımının ilk evrede ret edilmesine karar vermek için kullanılan kritik



ŞEKİL 2: Kepner/Chang'ın iki-evreli faz II klinik deneme tasarımı adımları.

- a_1 : 1. evrede reddedilmesi için ilaca cevap veren maksimum hasta sayısı
 r_1 : 1. evreden sonra ilacın etkin olduğuna ve bir sonraki çalışmaya geçmek için ilaca cevap verenlerin sayısının alt sınırı
 n_1 : 1. evredeki incelenen hasta sayısı
 c_1 : 1. evredeki cevap veren sayısı
 r : Deneme sonunda ilacın etkin olup olmadığına karar vermek için ilaca cevap veren maksimum hasta sayısı
 n : Denemede incelenen toplam hasta sayısı
 c : Deneme sonundaki cevap veren sayısı

cevap sayısıdır a_1 'dir (Şekil 2). Tip 2 tasarım için ise durum Tip 1 tasarım ile benzerdir. Tip 3 tasarımı bulmaya çalışan algoritma tip 1 ve tip 2 tasarımlarını bulmaya çalışan yöntemlerin birleşimidir. Örneklem büyüklüğü n_1 ve n_2 olan 1. ve 2. evredeki objektif cevaplayıcıların sayısı c_1 ve c_2 ile ifade edilsin. E, $c_1 \geq r_1 + 1$ ya da $[c_1 \leq r_1, c_2 \geq r_2 + 1]$ ($r_1 + 1 \leq n_1, r_2 + 1 \leq n$ ve $r_1 \leq r_2$ olmak üzere) şartlarını sağlayan bir olay olsun. İstenen tasarım, $P(E | p_0) \leq \alpha$ ve $P(E | p_0 + \Delta) \geq 1 - \beta$ olasılık şartlarında (n_1, r_1, n_2, r_2) dörtlüsünden oluşur. Başlangıç değerleri $n_1 = n_2 = 1$ olan $P(E | p_0)$ ve $P(E | p_0 + \Delta)$ olasılıklar, $r_1 \leq r_2$ 'nin $r_1 + 1 \leq n_1, r_2 + 1 \leq n$ ve $r_1 \leq r_2$ şartlarındaki tüm kombinasyonlar kullanılarak hesaplanır. Koşullar sağlanmadıkça n_2 bir artırılır ve araştırma r_1 ve r_2 'nin kombinasyonları, $n_1 = n_2$ oluncaya kadar devam eder. Hala geçerli bir koşullar sağlanmazsa n_1 bir artırılır ve $n_2 = 1$ yapılarak işlem devam eder. $p_0, \Delta > 0, \alpha, 1 - \beta$ parametrelerinin girdilerine ve tasarım tipine bağlı olarak yürütülen tasarımlarda toplam örneklem büyüklüğü N , tek-evreli test için gereken örneklem büyüklüğü n 'den daha küçük olabilir. Belirli giriş parametreleri için, verilerin k kez incelenmesi, kullanıcının belirlediği büyüklük ve güç kısıtlarının gerçekleştirilmesi ve bazen aynı giriş parametrelili düzgün en güçlü tek-evreli testine göre kullanılan hasta sayısından daha az hasta yer alması mümkündür. Bunun sebebi, grup ardaşık testin örnek uzayının, tek-evre testin örnek uza-

yına göre daha fazla örneklem noktası içermesidir. Anlamlılık seviyesi ve güç özelliklerinin kullanıcı tanımlı seviyelere yakın olan grup ardaşık tasarımlar bulunmasını mümkün kılar. Bir başka kayda değer sonuç ise, düzgün en güçlü tek-evreli testinin büyüklüğü ve gücü, p_0 ve $\Delta > 0$ girdileriyle kullanılırsa, herhangi üç tipe ait düzgün en güçlü tek-evreli testle aynı büyüklük ve güçte $p_0 + \Delta$ k-evreli tasarım üretebilir.

SOUTHWEST ONKOLOJİ GRUBU STANDART TASARIMI

Southwest Onkoloji Grubu (SWOG) standart yaklaşımı, iki-evreli tasarımları yaklaşık 0.05 düzeyinde, 0.90 güç ve birinci evrede H_1 hipotezi 0.02 düzeyinde reddedilmesiyle erken sonlandırıldığı durumda düşünülmektedir.¹² SWOG'un faz II tasarımına yaklaşımı denemelerin erken sonlandırılma ihtiyacı gibi çeşitli pratik nedenlere cevap olarak yavaş yavaş gelişmektedir. Çok merkezli denemelerin gerçekleri aşağıda belirtilmiştir.¹²

1. Çok merkezli çalışmalar, belirlenen sayıda hastayı çalışmaya aldıktan sonra sonlandırılmayaabilir. Çünkü bazı merkezlerin, denemeye katılmaları için önceden belirledikleri hastaları çalışmaya almaya devam etmelerinden dolayı belirli bir zaman periyodunun geçmesi gerekmektedir. Hasta sayısının artışının kesin kontrolü olmadan optimallik sağlanamaz. İlk evre sonucunun olasılığı p_1 altında, β 'dan küçük ise ilk evrede deneme durdurulur. Aksi takdirde anlamlılığı $\leq \alpha$ ve gücü $\geq 1 - \beta$ olan en küçük örneklem büyüklüğüne sahip tasarım seçilmiş olur. SWOG karar kuralları, elde edilen örneklem büyüklüğünün planlanan örneklem büyüklüğüne eşit değil ise kolayca uygulanabilir.

2. Dikkatli bir şekilde tanımlanan seçicilik kriterine rağmen önemli derecede hasta heterojenliği mevcuttur. Aktif olmayan bir ilacın hammaddesinin yer aldığı denemeyi erken sonlandırmak önemliyken, aktif bir ajanın kötü şansa sahip hastalardaki testlerine dayanarak reddedilmesini korumaya karşı katı olması da önemlidir. Bu konu p 'nin bazı seçenekleri altında beklenen örneklem büyüklüğünün minimizasyonu prensiplerinden daha zorlayıcıdır. 0.02 düzeyinde $H_1: p = p_1$ test edilir, çünkü durdurma olasılığını sıfır hipotezi altında önemsizleştirilmediğinde tutucu bulunmaktadır.

3. Dağılım kesikli ve örneklem büyüklüğü küçük olduğunda 0.05 anlamlılık düzeyi düşük olabilir. İkinci evrede sınır belirlemek ve anlamlılık düzeyini 0.05'e yakın elde etmek için tek-evrede 0.055 anlamlılık düzeyi kullanılmaktadır. Buna ek olarak deneme, H_0 'ın lehinde erken sonlandırılırsa anlamlılık düzeyi tek-evreli tasarımlara göre az da olsa azalmaktadır.

4. İlk evredeki toplam 0.05 anlamlılık, 0.90 güç ve 0.02 kuralı seçenekleri, birçok klinik araştırmacıların sezgileriyle yaklaşık olarak uyum gösterdiği ilk evre, planlanan toplamın, N , $1/2$ 'si ve $p_0(1-p_0)$, $p_1(1-p_1)$ yakın olduğu zamanki sürecin sonucudur. Bu durumda, cevap olasılığının tahmini, yaklaşık olarak p_0 'dan az olduğu zaman ilk evrede durdurma gerçekleşir.

X_1 ilk evredeki cevap sayısı ise ilk evrede H_0 'ı kabul etmek için karar sınırları, r_1 , yaklaşık olarak $B(X_1 \leq r_1 | p_1, N_1) = 0.02$ sağlar ve böylece

$$r_1 / N_1 \approx p_1 - \frac{(p_1 - p_0)2.05\sqrt{2}}{1.65\{p_0(1-p_0) / p_1(1-p_1)\}^{1/2} + 1.28}$$

$p_0(1-p_0) = p_1(1-p_1)$ olduğu zaman $0.99p_0 + 0.01p_1$ 'dir.

İkinci evrede, eğer cevap olasılığının tahmini yaklaşık olarak $(p_1 + p_0)/2$ 'den büyükse ajanın bir sonraki çalışmayı garantilediği sonucuna varılır. Eğer X , iki evredeki toplam cevap sayısı ise ikinci evreden sonra H_0 'ı reddetmek için sınır olan r , yaklaşık olarak

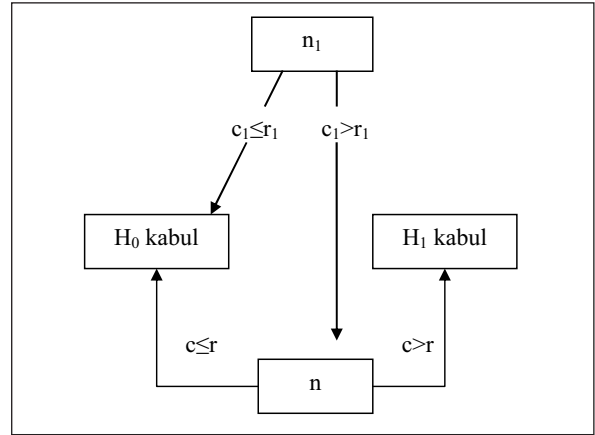
$$r / N \approx p_0 - \frac{(p_1 - p_0)}{1 + 1.28\{p_1(1-p_1)\}^{1/2} / 1.65\{p_0(1-p_0)\}^{1/2}}$$

$p_0(1-p_0) = p_1(1-p_1)$ olduğu zaman $0.44p_0 + 0.56p_1$ 'dir.

SWOG tasarımının çalışma şekli Şekil 3'te gösterilmiştir.

FAZ II KLİNİK DENEMELERİ İÇİN FARKLI DÜZEYLERDEKİ (α, β) VE (p_1, p_0) İÇİN ÖRNEKLEM BÜYÜKLÜKLERİ

Literatürlerde faz II klinik çalışmalarının evre sayısı hakkında farklı yorumlar bulunmaktadır. Fakat çoğunluk tarafından evre sayısının 2 olmasının yeterli olduğu belirtilmektedir. İki-evreli faz II klinik denemelerden Simon'un geliştirdiği optimal ve minimaks tasarım, Kepner/Chang ve SWOG tasarımları (α, β) 'nin (0.05, 0.10) olduğu, $p_1 - p_0$ farkının



ŞEKİL 3: SWOG iki-evreli faz II klinik deneme tasarım adımları.

r_1 : Denemenin sonlandırılması için ilaca cevap veren maksimum hasta sayısı

n_1 : 1.evredeki incelenen hasta sayısı

c_1 : 1.evrede gözlenen cevap veren sayısı

r : Deneme sonunda ilacın etkin olup olmadığına karar vermek için ilaca cevap veren hasta sayılarının karşılaştırıldığı hasta sayısı

n : Denemede incelenen toplam hasta sayısı

c : 2. evre sonunda gözlenen cevap veren sayısı

0.10, 0.15 ve 0.20 olduğu durumlarda r_1 , n_1 , r_1/n_1 , r , n ve r/n parametreleri için incelenmiştir. Sonuçlar tablolarda belirtilmiştir (Tablo 1-3).

Hesaplamalarda NCSS-PASS istatistiksel paket programından ve American Cancer Society (www.cryptnet.net/kepner) ve Southwest Oncology Group (<http://www.swogstat.org/stat/public/Help/twostage.html>) web sayfalarından yararlanılmıştır.

İki-evreli faz II klinik denemelerinde geliştirilen tasarımlar genelde birbirine benzemekle birlikte tamamıyla aynı değildir. Simon'un optimal, minimaks tasarımları ve SWOG tasarımı için ilk evre sonunda H_0 hipotezinin reddedilme olasılığı mümkün değil iken, Kepner/Chang tasarımında ilk evre sonunda yeterli cevap sayısı elde edilmesi durumunda H_0 hipotezi reddedilebilir. Simon'a göre iki-evreli faz II klinik deneme tasarımlarında ilk evreden sonra ilacın etkin olduğuna karar vermek etik değildir.⁷ Ancak, hasta sayısının kısıtlı olduğu ve ilacın pahalı olduğu durumlar için 1. evreden sonra ilacın etkin olduğuna, bir başka deyişle H_0 hipotezin erken reddine karar verilebilir.⁷

İki-evreli faz II klinik denemeleri için geliştirilen tasarımların çeşitli $p_1 - p_0$ farkları, α ve β değer-

TABLO 1: $p_1-p_0=0.10$ ve $\alpha=0.05$, $\beta=0.10$ için tasarımların karşılaştırması.

p_0	p_1	Tasarım	r_1	n_1	r_1/n_1	r	n	r/n	α	β
0,05	0,15	Optimal	2	37	0,05	7	84	0,08	0,048	0,099
		Minimaks	2	46	0,04	7	77	0,09	0,037	0,099
		Kepner/Chang	5	38	0,13	7	76	0,09	0,040	0,097
		SWOG	1	40	0,03	7	75	0,09	0,033	0,111
0,10	0,20	Optimal	4	42	0,10	17	121	0,14	0,049	0,099
		Minimaks	6	70	0,09	16	109	0,15	0,043	0,100
		Kepner/Chang	11	55	0,20	16	109	0,15	0,045	0,100
		SWOG	4	55	0,07	16	110	0,15	0,046	0,094
0,15	0,25	Optimal	9	57	0,16	31	162	0,19	0,050	0,099
		Minimaks	11	78	0,14	27	136	0,20	0,048	0,100
		Kepner/Chang	18	68	0,26	27	136	0,20	0,049	0,100
		SWOG	9	70	0,13	28	140	0,20	0,042	0,103
0,20	0,30	Optimal	16	76	0,21	43	175	0,24	0,050	0,099
		Minimaks	18	92	0,20	40	160	0,25	0,049	0,100
		Kepner/Chang	26	80	0,33	40	160	0,25	0,050	0,098
		SWOG	15	80	0,19	40	160	0,25	0,049	0,101
0,25	0,35	Optimal	22	82	0,27	64	218	0,29	0,048	0,100
		Minimaks	30	125	0,24	54	179	0,30	0,048	0,100
		Kepner/Chang	33	90	0,37	54	179	0,30	0,050	0,100
		SWOG	21	90	0,23	54	180	0,30	0,052	0,094
0,30	0,40	Optimal	29	91	0,32	79	229	0,34	0,050	0,099
		Minimaks	41	142	0,29	68	193	0,35	0,050	0,100
		Kepner/Chang	40	101	0,40	68	194	0,35	0,050	0,099
		SWOG	29	100	0,29	69	195	0,35	0,044	0,110
0,35	0,45	Optimal	32	88	0,36	94	237	0,40	0,050	0,099
		Minimaks	44	127	0,35	83	206	0,40	0,049	0,100
		Kepner/Chang	50	105	0,48	82	205	0,40	0,050	0,098
		SWOG	37	105	0,35	81	205	0,40	0,093	0,063
0,40	0,50	Optimal	39	94	0,41	107	239	0,45	0,050	0,100
		Minimaks	76	176	0,43	96	212	0,45	0,050	0,100
		Kepner/Chang	66	112	0,59	96	215	0,45	0,049	0,098
		SWOG	47	110	0,43	96	215	0,45	0,073	0,101

leri için sonuçları tablolarda verilmiştir (Tablo 1-3). Faz II denemeleri için bu iki hata tipi önemlidir. $\beta < \alpha$ olması istenmeyen bir durumdur ve örneğin $\alpha = \beta = 0.05$ 'e dayanan bir tasarım faz II denemelerin bir çoğunda büyük örneklem büyüklüklerini içerir.⁷

Simon, tümöre karşı geliştirilen yeni bir ilaç için (p_0, p_1) değerleri (0.05, 0.20), (0.05, 0.25) veya (0.10, 0.25) olan tasarımların daha uygun olacağını belirtmiştir.⁷ Bunun sebebi, birçok yeni ilaç ortak tümörlere karşı genelde etkisizdir. Ayrıca dirençli hastalıklara karşı göreceli olarak (%20-25) oranında ılımlı cevap oranı sağlayan yeni ilaçlar, sonraki ge-

lişmeleri ilgilendirir.⁷ Simon'a göre, pilot çalışmalar için istenen ortak fark derecesi $p_1 - p_0 = 0.20$ düzeyindedir.⁷ Büyük farkları ayırmak için düzenlenen deneme tasarımları gerçekçi ve bilgilendirici değildir. Daha küçük farklar için çok büyük örneklem büyüklükleri gerekmesinden dolayı Simon'a göre $p_1 - p_0 = 0.15$ farkı en küçük fark olarak gösterilebilir.⁷ Ayrıca kontrol eksikliği küçük farkları ayırmaya dayanan denemelerin yorumlanması kısıtlamaktadır. Ancak, literatürlerde yer alan diğer faz II deneme tasarımları için $p_1 - p_0 = 0.10$ farkı içinde inceleme gerçekleştirilmiştir.

TABLO 2: $p_1-p_0=0.15$ ve $\alpha=0.05$, $\beta=0.10$ için tasarımların karşılaştırması.

p_0	p_1	Tasarım	r_1	n_1	r_1/n_1	r	n	r/n	α	β
0,05	0,20	Optimal	1	21	0,05	4	41	0,10	0,046	0,098
		Minimaks	1	29	0,03	4	38	0,11	0,039	0,100
		Kepner/Chang	3	20	0,15	4	38	0,11	0,044	0,098
		SWOG	0	20	0,00	4	40	0,10	0,047	0,080
0,10	0,25	Optimal	3	28	0,11	9	57	0,16	0,048	0,099
		Minimaks	3	31	0,10	9	55	0,16	0,042	0,099
		Kepner/Chang	6	27	0,22	9	54	0,17	0,046	0,098
		SWOG	2	30	0,07	9	55	0,16	0,044	0,091
0,15	0,30	Optimal	5	30	0,17	17	82	0,21	0,046	0,099
		Minimaks	6	42	0,14	14	64	0,22	0,048	0,100
		Kepner/Chang	10	32	0,31	14	64	0,22	0,050	0,098
		SWOG	4	35	0,11	15	65	0,23	0,028	0,140
0,20	0,35	Optimal	8	37	0,22	22	83	0,27	0,049	0,099
		Minimaks	8	42	0,19	21	77	0,27	0,044	0,100
		Kepner/Chang	14	39	0,36	21	77	0,27	0,046	0,098
		SWOG	7	40	0,18	21	75	0,28	0,034	0,126
0,25	0,40	Optimal	10	37	0,27	31	99	0,31	0,049	0,100
		Minimaks	13	57	0,23	27	83	0,33	0,047	0,100
		Kepner/Chang	17	42	0,40	27	83	0,33	0,049	0,100
		SWOG	10	45	0,22	28	85	0,33	0,038	0,112
0,30	0,45	Optimal	13	40	0,33	40	110	0,36	0,048	0,099
		Minimaks	27	77	0,35	33	88	0,38	0,050	0,099
		Kepner/Chang	20	45	0,44	34	90	0,38	0,050	0,099
		SWOG	12	45	0,27	34	90	0,38	0,044	0,103
0,35	0,50	Optimal	16	43	0,37	44	105	0,42	0,049	0,100
		Minimaks	16	46	0,35	40	94	0,43	0,049	0,100
		Kepner/Chang	25	48	0,52	41	96	0,43	0,048	0,095
		SWOG	17	50	0,34	41	95	0,43	0,039	0,113
0,40	0,55	Optimal	19	45	0,42	49	104	0,47	0,050	0,100
		Minimaks	24	62	0,39	45	94	0,48	0,049	0,100
		Kepner/Chang	28	47	0,60	45	94	0,48	0,049	0,100
		SWOG	19	50	0,38	46	95	0,48	0,038	0,120

Simon'ın her iki tasarımında ve SWOG tasarımında erken sonlandırma olasılığı varken Kepner/Chang tasarımında bu durum söz konusu değildir. Kepner/Chang tasarımında ise denemenin başında ilacın reddedilmesi veya kabul edilmesi mümkündür.

Faz II klinik deneme tasarımlarından optimal, minimaks ve Kepner/Chang tasarımları α ve β hata oranları bakımından birbirlerinden çok farklılık göstermemektedir. Fakat SWOG tasarımı için ise bu iki hata oranı çok büyük farklılıklar göstermekte-

dir. Bununla birlikte Kepner/Chang tasarımının 1. evre sonundaki cevap oranı diğer tasarımlarinkinden daha yüksektir. Bu durum ilacın etkin olduğuna karar verilmesini olumsuz yönde etkilemektedir. SWOG tasarımı için 1.evre sonundaki cevap oranı en düşüktür. Deneme sonundaki cevap oranlarına bakıldığında en yüksek oran SWOG tasarımına ait iken optimal tasarım için bu oran diğer tasarımlardan daha düşüktür. Bu sebeple optimal tasarımın seçilmesi uygun olabilir. Ancak optimal tasarım için gereken örneklem büyüklüğü

TABLO 3: $p_1-p_0=0.20$ ve $\alpha=0.05$, $\beta=0.10$ için tasarımların karşılaştırması.

p_0	p_1	Tasarım	r_1	n_1	r_1/n_1	r	n	r/n	α	β
0,05	0,25	Optimal	0	9	0,00	3	30	0,10	0,049	0,098
		Minimaks	0	15	0,00	3	25	0,12	0,034	0,099
		Kepner/Chang	2	13	0,15	3	25	0,12	0,045	0,097
		SWOG	0	15	0,00	4	25	0,16	0,034	0,099
0,10	0,30	Optimal	2	18	0,11	6	35	0,17	0,047	0,098
		Minimaks	2	22	0,09	6	33	0,18	0,041	0,098
		Kepner/Chang	5	17	0,29	6	33	0,18	0,042	0,094
		SWOG	1	20	0,05	8	35	0,23	0,020	0,134
0,15	0,35	Optimal	3	19	0,16	10	44	0,23	0,048	0,095
		Minimaks	3	23	0,13	9	38	0,24	0,048	0,099
		Kepner/Chang	7	19	0,37	9	38	0,24	0,050	0,096
		SWOG	2	20	0,10	10	40	0,25	0,030	0,124
0,20	0,40	Optimal	4	19	0,21	15	54	0,28	0,048	0,096
		Minimaks	5	24	0,21	13	45	0,29	0,048	0,100
		Kepner/Chang	9	24	0,38	13	44	0,30	0,048	0,099
		SWOG	4	25	0,16	13	45	0,29	0,052	0,086
0,25	0,45	Optimal	6	22	0,27	19	57	0,33	0,047	0,100
		Minimaks	6	26	0,23	17	49	0,35	0,045	0,100
		Kepner/Chang	11	25	0,44	17	49	0,35	0,049	0,094
		SWOG	5	25	0,20	18	50	0,36	0,029	0,129
0,30	0,50	Optimal	8	24	0,33	24	63	0,38	0,050	0,097
		Minimaks	7	24	0,29	21	53	0,40	0,047	0,098
		Kepner/Chang	14	27	0,52	21	53	0,40	0,049	0,093
		SWOG	8	30	0,27	22	55	0,40	0,041	0,090
0,35	0,55	Optimal	7	20	0,35	26	59	0,44	0,049	0,099
		Minimaks	12	37	0,32	24	53	0,45	0,045	0,100
		Kepner/Chang	15	27	0,56	24	53	0,45	0,048	0,099
		SWOG	10	30	0,33	25	55	0,45	0,040	0,102

diğer tasarımlardan oldukça fazladır. Minimaks tasarımı, α , β hata oranları, 1.evre sonunda ve deneme sonundaki cevap oranları bakımından optimal tasarım ile çok farklılık göstermemektedir. Minimaks tasarım, optimal tasarıma göre daha küçük örneklem büyüklüğü içerir. Ancak 1. evre sonundaki örneklem büyüklüğü optimal tasarımda minimaks tasarıma göre daha küçüktür. Bu yüzden çalışmada yer alan denekler heterojen bir grup oluşturuyorsa minimaks tasarımın kullanılmasında yarar vardır.

Hastalıklara karşı tedaviyi aramak veya hastalıkların ilerlemesini durdurmak tıp biliminin en büyük amaçlarından. Bu amaçla yapılan çalışmalarda etik unsurlara önem verildiği gibi ve ekonomik durumlar da kısıtlılıklar getirmektedir. Bu koşullarda klinik deneme tasarımlarında örneklem büyüklüğü oldukça önem göstermektedir. Etik ve ekonomik açıdan optimal noktayı belirleyip faydayı ortaya çıkartmak için klinik deneme dizaynlarının kullanılması gerekmektedir.

KAYNAKLAR

1. Mıstık S, Mazırcıoğlu M, Göğüsten B, Erenmemişoğlu A. [Bioequivalence trials and ethical applications in Turkey]. *Turkiye Klinikleri J Med Ethics* 2002;10(4):285-9.
2. Demir M, Örnek Büken N. [Comparison between the regulation of regarding clinical drug research and the regulation of regarding clinical trials]. *Turkiye Klinikleri J Med Ethics* 2009;17(3):168-75.
3. Chow SH, Liu JP. *Intrroduction. Design and Analysis of Clinical Trials. A Wiley-Interscience Publication.* 1st ed. New Jersey: John Wiley&Sons, INC; 1998. p.1-43.
4. Meinert CL. *Introduction. Clinical Trials: Design, Conduct, and Analysis.* 1st ed. New York: Oxford University Press; 1986. p.3-10.
5. Piantadosi S. *Preliminaries. Clinical Trials: A Methodological Perspective.* 1st ed. New York: A Wiley-Interscience Publication; 1997. p.1-9.
6. Spilker B. *Guide to Clinical Trials. Chapter 1.* 1st ed. New York: Raven Press; 1991. p. 4-15.
7. Simon R. *Optimal two-stage designs for phase II clinical trials.* *Control Clin Trials* 1989;10(1): 1-10.
8. Chang MN, Therneau TM, Wieand HS, Cha SS. *Designs for group sequential phase II clinical trials.* *Biometrics* 1987;43(4):865-74.
9. Herson J. *Predictive probability early termination plans for phase II clinical trials.* *Biometrics* 1979;35(4):775-83.
10. Bryant J, Day R. *Incorporating toxicity considerations into the design of two-stage phase II clinical trials.* *Biometrics* 1995;51(4):1372-83.
11. Shuster J. *Optimal two-stage designs for single arm phase II cancer trials.* *J Biopharm Stat* 2002;12(1):39-51.
12. Green SJ, Dahlberg S. *Planned versus attained design in phase II clinical trials.* *Stat Med* 1992;11(7):853-62.
13. Jung SH, Carey M, Kim KM. *Graphical search for two-stage designs for phase II clinical trials.* *Control Clin Trials* 2001;22(4):367-72.
14. Colton T, McPherson K. *Two-stage plans compared with fixed-sample-size and wald SPRT plans.* *J Am Stat Assoc* 1976;71(353): 80-6.
15. McPherson K. *On choosing the number of interim analyses in clinical trials.* *Stat Med* 1982; 1(1):25-36.
16. Jennison C. *Efficient group sequential tests with unpredictable group sizes.* *Biometrika* 1987;74(1):155-66.
17. Chen TT, Ng TH. *Optimal flexible designs in phase II clinical trials.* *Stat Med* 1998;17(20): 2301-12.
18. Kepner JL, Chang MN. *Samples of exact k-stage group sequential designs for Phase II and Pilot studies.* *Control Clin Trials* 2004; 25(3):326-33.