

# KLİNİK ARAŞTIRMALARDA ÖRNEK GENİŞLİĞİ BELİRLEME

## SAMPLE SIZE DETERMINATION FOR CLINICAL RESEARCH

**Dr. Seval Kul**

Gaziantep Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Gaziantep, Türkiye

**e-posta:** sevalkul@gantep.edu.tr

doi:10.5152/pb.2011.11

### Özet

Bir çalışmanın planlanma aşamasında yeterli büyüklükteki örneklem genişliğini tahmin etmek gerekir. Bunun birçok nedeni vardır. Öncelikli olarak güç analizi yardımıyla kaynakların ve emeklerin boşa harcanmasının önüne geçilir. Ayrıca gereğinden küçük örneklerle çalışmanın sonucu düşük güce sahip bir çalışma yapmak olacaktır. Gereğinden çok büyük örnek genişliği ile gerçekleştirilen çalışmalarda ise klinik olarak anlamlı olmayan iki uygulamanın istatistiksel olarak anlamlı çıkması mümkündür. Bu çalışmada örnek genişliği belirlemenin temel ilkelerinden bahsedilmiş ve iki hipotetik örnek için örnek genişliği belirlenmiştir.

**Anahtar kelimeler:** Güç analizi, örnek büyüklüğü

### Abstract

The required sample size for any intervention must be estimated at the planning stage of a study. There are several reasons for this. First, power analysis prevents waste of resources and efforts. In addition, too small a sample size results in the loss of power. Also, an unnecessarily large sample size may increase the probability of statistically significant results which are not clinically significant. In this study, the principles of sample size determination were mentioned and sample sizes were estimated for two hypothetical examples.

**Key words:** Power analysis, sample size

Klinik bir çalışmada hipotezler oluşturulurken öncelikle araştırmanın popülasyonu belirlenir ve ardından bu popülasyonu iyi temsil ettiği düşünülen bir örneklem ile çalışılır. Buradaki amaç elimizdeki örneklemden yararlanarak popülasyon hakkında bir takım kararlar almaktır (1).

Aslında her klinik çalışmada yapılan şey bir bilimsel hipotezi test etmektir. İstatistikte bir hipotezin doğruluğunu test etmek için hipotez takımı oluşturulur. Bu hipotez takımının iki bileşeni vardır (2). Birincisi Kontrol veya Yokluk hipotezi ( $H_0$ ) olarak bilinir. Bu hipotezde her zaman gruplar arasında fark olmadığı belirtilir. Yokluk hipotezinin içermediği diğer tüm durumları içeren hipoteze ise Alternatif veya Karşıt hipotez ( $H_1$ ) adı verilir (2). Hipotezler sözel olarak ifade edilebileceği gibi semboller ile de ifade edilebilirler. İki farklı tedavi alan hastaların hastaneden taburcu olma süreleri arasında anlamlı bir fark olup olmadığının karşılaştırılacağı

bir çalışma planladığımızı düşünelim. Böyle bir araştırmanın hipotez takımı aşağıdaki gibi olacaktır.

$$H_0: \mu_A = \mu_B$$

$$H_1: \mu_A \neq \mu_B$$

$H_0$  hipotezinde belirtilen ortalama taburcu olma süreleri bakımından A ve B tedavileri arasında anlamlı bir farklılık olmadığıdır. Bunun alternatifi olan karşıt hipotezde ise anlamlı farklılık olduğu belirtilmektedir.

Bir çalışmayı planlama aşamasında araştırmacının aklındaki en önemli sorulardan biri hem klinik hem de istatistiksel olarak anlamlı bir sonuç bulabilmek için en az kaç kişi ile çalışması gerektiğidir. Bu sayının belirlenmesine etkisi olan birçok faktör vardır (3).

Örnek genişliği belirlenmesine etkisi olan faktörler:

**1. Sonuç değişkenin tipi:** Bir araştırmanın sonucunda karşınıza en yaygın olarak çıkabilecek 2 farklı çeşit sonuç değişkeni olabilir. Bunlar; bir hastalığın, yan etkinin ya da iyileşmenin görülme oranı, yüzdesi gibi ifade

edilen kategorik değişkenler ya da bir ölçme aracı yardımıyla elde edilmiş vücut fonksiyonlarını belirtmek için kullanılan sayısal olarak ifade edilen değişkenler değişkenlerdir. Sonuç değişkeninin tipine göre uygulanacak istatistiksel analizler farklılık göstereceği için örneklem genişliği belirlenme şeklide farklıdır.

**2. Deneme Planı:** Çalışmada nasıl bir deneme planının kullanılacağı da gerekli minimum örnek genişliğini belirlemede etkilidir. Çünkü yapılan güç analizi sonucunda hesaplanan sayı her grup için gerekli olan minimum denek sayısıdır. Bazı durumlarda aynı kişilerden belirli bir uygulamadan önce ve sonra ölçümler alınması planlanır. Bu durumda bir tek çalışma grubunun farklı zamanlardaki değerleri karşılaştırılacağı için belirlenen örnek genişliği kadar kişi ile çalışmak yeterli olacaktır. Diğer taraftan eğer bir çalışmada 2 farklı grup olacaksa örneğin, yeni bir yöntem karşı standart tedavinin karşılaştırılması hedefleniyorsa çalışmaya dahil edilen kişi sayısı 2 katına çıkacaktır.

**3. I. Tip hata miktarı:** Hipotez testi sırasında araştırmacının yapabileceği hatalardan biridir. Aslında iki uygulama arasında fark yokken araştırma sonucunda fark bulunabilmesi ihtimalidir (3).  $\alpha$  hatası ya da p değeri olarak da bilinir (4). Örnek genişliği belirlenirken genellikle I. Tip hata miktarı 0.05 ya da 0.01 olarak kabul edilir (4).

**4. II. Tip hata:** Hipotez testi sırasında araştırmacının yapabileceği bir diğer hata ise II. tip hatadır (3). Aslında iki uygulama arasında fark var iken, araştırma sonucunda farkın bulunamaması ihtimalidir.  $\beta$  hatası olarak da bilinir. Örnek genişliği belirlenirken genellikle II. Tip hata miktarı 0.20 olarak kabul edilir (4).

**5. Etki Büyüklüğü:** Klinik olarak anlamlı farklılığın ortaya konması için ilgilenilen sonuç değişkenine göre iki ortalama ya da iki oran arasındaki beklenen farklılıktır. Etki büyüklüğüne literatür tarayarak veya hekimin geçmiş deneyimleri ile karar vermesi mümkün iken herhangi bir şekilde örneği olmayan bir çalışma için bir pilot çalışma düzenlenerek belirlenebilir (3). Gruplar arasındaki beklenen farklılık yani etki büyüklüğü ne kadar büyük olursa o farkı ortaya koymak daha kolay olacağı için daha küçük örnek genişliği ile çalışmak yeterli olur.

### Güç analizi

Testin gücü, bir testin gerçekte var olan farkı bulabilme yeteneğidir. Bu nedenle II. Tip hata miktarı ile ilişkilidir (5). II. Tip hatanın yani var olan bir farkın bulunamama olasılığının 0.20 olduğu durumda, yapılacak hipotez testinde hedeflenen testin gücü 0.80 olacaktır. Örnek genişliği belirlemeye güç analizi denmesinin

nedeni örnek genişliğini çalışmanın başında belirleyerek çalışacağımız gücü kontrol altına almamızdır (6).

Örnek genişliğin formüller yardımıyla hesaplanması oldukça kolaydır (6). Bunun yanı sıra günümüzde birçok ücretsiz paket programa da internet aracılığı ile ulaşmak mümkündür. Bu çalışmada uygulamalar PS version 3.0 paket programında gösterilmiştir (7).

### Kategorik değişkenler için örnek genişliği belirleme

Sonuç değişkeninin bir hastalığın, yan etkinin ya da iyileşmenin görülme oranı şeklinde olduğu çalışmalarda amaç bu oranlar arasında anlamlı bir farklılık olup olmadığının test edilmesidir. Eğer ikiden fazla oranın karşılaştırılması hedefleniyorsa örnek genişliği birbirine en yakın oran beklentisi olan yani etki büyüklüğü en küçük olan iki grup için hesaplanmalıdır (8).

### Uygulama 1

Koah ve Astım hastalarında kalp yetmezliği görülme oranları bakımından anlamlı bir farklılığın olup olmadığının test edilmek istendiği bir çalışma planladığımızı varsayalım. Koah grubunda beklenen kalp yetmezliği oranı 0.60 ve astım grubunda 0.40 olduğu durumda I. Tip hata miktarı 0.05, testin gücü 0.80 iken her iki grupta gerekli minimum örnek genişliği kaç olmalıdır?

Program tıkladığında aşağıdaki ana ekran gelmektedir. Sonuç değişkenimiz kategorik olduğu için Dichotomus seçeneği seçilir (Şekil 1).

Çalışmanın dizaynını belirtmek için Desing yazan alandan gerekli seçimler yapılmalıdır. Astım ve Koah grubundaki bireyler birbirinden bağımsız iki grup oldukları için Independent seçeneği işaretlenir. Alternatif hipotezde iki oran karşılaştırılacaktır ve analiz içinde klasik ki-kare testi seçilebilir. İki bağımsız oranın Ki-kare analizi ile karşılaştırılmasının planladığı bir çalışmada ileriye ya da geriye dönük çalışma arasında örnek genişliği bakımından herhangi bir farklılık olmayacaktır. Eğer risk faktörlerinin belirlenmesine yönelik bir çalışma ise çalışma geriye dönük planlanacaksa vaka kontrol çalışması ve odds oranı, ileriye dönük planlanacaksa prospective seçeneği ve relative risk seçilmelidir.

Input kısmında I. Tip hata, hedeflenen güç ve iki grupta beklenen oranlar girilmelidir. Burada m ile belirtilen ise kaç tane kontrol grubu olduğudur. Calculate tıklandıktan sonra Output kısmında her grup için gerekli minimum örnek genişliği hesaplanacaktır. Description kısmında ise detaylı olarak yapılan işlem açıklanmaktadır. Bu açıklamalar kopyalanıp yapııştırılabilir özelliktedir.

Bu durumda I. Tip hata miktarı 0.05, testin gücü 0.80 iken iki oran arasında anlamlı bir fark bulabilmemiz için her grupta gerekli olan minimum örnek genişliği 97 olarak belirlenir.

**Power and Sample Size Program: Main Window**

File Edit Log Help

Survival | t-test | Regression 1 | Regression 2 | Dichotomous | Mantel-Haenszel | Log

Output: [Studies that are analyzed by chi-square or Fisher's exact test](#)

What do you want to know? Sample size

Case sample size for uncorrected chi-squared test: 97

Design:

- Matched or independent? Independent
- Case control? Case-Control
- How is the alternative hypothesis expressed? Two proportions
- Uncorrected chi-square or Fisher's exact test? Uncorrected chi-square test

Input:

$\alpha$ : 0.05  $p_0$ : 0.60

power: 0.80  $p_1$ : 0.40

$m$ : 1

Buttons: Calculate, Graphs

Description:

We are planning a study of independent cases and controls with 1 control(s) per case. Prior data indicate that the probability of exposure among controls is 0.6. If the true probability of exposure among cases is 0.4, we will need to study 97 case patients and 97 control patients to be able to reject the null hypothesis that the exposure rates for case and controls are equal with probability (power) 0.8. The Type I error probability associated with this test of this null hypothesis is 0.05. We will use an uncorrected chi-squared statistic to evaluate this null hypothesis.

Şekil 1. Kategorik değişkenler için örnek genişliği belirleme

**Power and Sample Size Program: Main Window**

File Edit Log Help

Survival | t-test | Regression 1 | Regression 2 | Dichotomous | Mantel-Haenszel | Log

Output: [Studies that are analyzed by t-tests](#)

What do you want to know? Sample size

Sample Size: 18

Design:

- Paired or independent? Paired

Input:

$\alpha$ : 0.05  $\delta$ : 0.56

power: 0.80  $\sigma$ : 0.8

Buttons: Calculate, Graphs

Description:

We are planning a study of a continuous response variable from matched pairs of study subjects. Prior data indicate that the difference in the response of matched pairs is normally distributed with standard deviation 0.8. If the true difference in the mean response of matched pairs is 0.56, we will need to study 18 pairs of subjects to be able to reject the null hypothesis that this response difference is zero with probability (power) 0.8. The Type I error probability associated with this test of this null hypothesis is 0.05.

Şekil 2. Sayısal değişkenler için örnek genişliği belirleme

## Sayısal Değişkenler İçin Örnek Genişliği Belirleme

Bir ölçme aracı yardımıyla elde edilmiş vücut fonksiyonlarını belirtmek için kullanılan sayısal olarak ifade edilen değişkenlerin sonuç değişkeni olduğu çalışmalarda amaç gruplar arasında ortalamalar bakımından anlamlı bir farklılık olup olmadığının test edilmesidir.

Etki büyüklüğünün hesaplanabilmesi için gruplarda ilgilenilen ölçüm bakımından klinik anlamlı bir farklılık oluşması için ortalamalar arasındaki farkın en az kaç olduğunun ve beklenen standart sapma değerinin tahmin edilmesi gereklidir. Bu değerler daha önceden yapılmış çalışmalardan elde edilebileceği gibi bir pilot çalışmadan da hesaplanabilirler.

### Uygulama 2

Astım hastalarında uygulanan yeni bir tedavi sonucunda FVC değerlerinde klinik olarak anlamlı bir artış beklenmektedir. Daha önceden yapılmış çalışmalarda astım hastaları için FVC değerinin ortalamasının  $2.8 \pm 0.8$  olduğu bilinmektedir. Bu değer kullanılarak yeni ilaç ile %20 arttırılacağı düşünülmüştür. Aynı bireylerin hem tedaviden önce hem de tedaviden sonraki FVC değerleri karşılaştırıldığında ölçümler arasında anlamlı fark bulabilmek için  $\alpha = 0.05$ ,  $1 - \beta = 0.80$  iken her çalışma grubunda en az kaç birey olmalıdır?

Program tıkladığında aşağıdaki ana ekran gelmektedir. Sonuç değişkenimiz sayısal olduğu için t test seçeneği seçilir (Şekil 2). Çalışmanın dizaynını belirtmek için Desing yazan alandan gerekli seçimler yapılmalıdır. Aynı bireylerden tedavi öncesi ve sonrası ölçümler alınacağı için ölçümler birbiri ile bağımlı olacaktır bu nedenle paired seçeneği işaretlenir. Input kısmında I. Tip hata (0.05), hedeflenen güç (0.80), ortalamalar arasındaki beklenen fark ( $\delta = 2.8 * 0.20 = 5.6$  lık artış) ve standart sapma değeri ( $\sigma = 0.8$ ) girilmelidir. Calculate tıklandıktan sonra Output kısmında her grup için gerekli minimum örnek genişliği hesaplanacaktır. Description kısmında ise detaylı olarak yapılan işlem açıklanmaktadır. Bu açıklamalar kopyalanıp yapıştırılabilir özelliktedir. Bu durumda I. Tip hata miktarı 0.05, testin gücü 0.80 iken iki ortalama arasında anlamlı bir fark bulabilmemiz için çalışma grubunda gerekli olan minimum örnek genişliği 18 olarak belirlenir.

### Kaynaklar

1. Simon SD. Statistical Evidence in Medical Trials, What Do the Data Really Tell Us? Oxford: Oxford press: 2006; 138-41.
2. Sümbüloğlu K, Sümbüloğlu V, Biyoistatistik, 11.baskı, Ankara: 2005; 49-52.

3. Smoller SW. Biostatistics and epidemiology, a primer for health and biomedical professionals, Newyork: Springer; 2003; 153-7.
4. Dawson B, Trap RG. Basic and clinical biostatistics, Lange Medical Books/McGraw-Hill, 2004; 94-5.
5. Moyé LA. Statistical reasoning in medicine, the intuitive p-value primer. Newyork: Springer; 2006; 141-3.
6. Le CT. Introductory Biostatistics, New Jersey: Wiley İnterscience; 2003; 451-4.
7. <http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize>
8. Dell RB, Holleran S, Ramakrishnan R. Sample Size Determination, ILAR Journal 2002; 43: 208-13.