

**DERLEME**

**İSTATİSTİKSEL GÜÇ ANALİZİ VE HEMŞİRELİK ARAŞTIRMALARINDA  
KULLANIMI: TEMEL BİLGİLER**

**Cantürk ÇAPIK\***

Alınış Tarihi: 30.05.2014

Kabul Tarihi: 29.09.2014

**ÖZET**

*En yalın tanımı ile istatistiksel güç, analizlerde anlamlı bir fark saptanması durumunda bu farkın gerçek olabilme ihtimalidir. İstatistiksel güç, örneklem sayısından etkilendiği için araştırmacılar örneklem sayısını güç analizi ile hesaplayabilmektedirler. Birçok akademik derginin çalışmanın gücünü istemesi, etik komitelerin örneklem sayısı hakkında açıklama istemesi ve araştırmacıların daha kesin bulgular sunmak istemesi gibi faktörler güç analizinin daha fazla dikkate alınmasını gerektirmektedir. Bir çalışmanın gücünü etkileyen birçok faktör olmasına rağmen Tip I hata, etki büyüklüğü ve örneklem büyüklüğü bunlardan en önemlileridir. İstatistiksel güç hesabı önemli olmasına rağmen, günümüzde hemşirelik araştırmalarında güç analizi yaparak örneklem sayısını belirlemenin yaygın olduğunu söylemek güçtür. Bu makale güç analizi hakkında giriş bilgisi sağlamaktadır. Makalede, güç analizi ile ilgili bazı istatistiksel terimlerin aktarılmasından sonra, temel bilgilerin, güç analizi çeşitlerinin ve güç analizinin örneklem hesabında nasıl kullanılacağına yönelik bilgiler verilmiştir. Son olarak da bazı hemşirelik çalışmalarından örnekler sunulmuştur.*

**Anahtar Kelimeler:** Güç analizi; hemşirelik; araştırma; istatistik.

**ABSTRACT**

**Statistical Power Analysis and It's Use in Nursing Studies: Basic Information**

*With its simplest definition, the statistical power is the possibility for a significant difference between analysis. Since the statistical power is affected by the number of samples, researchers may calculate the number of samples with the power analysis. Factors such as the requirement of many academic journals for the study's power, as well as the requirement of ethics committees for an explanation about the number of samples and the requirement of researchers to represent more precise findings enable the power analysis must be taken into consideration more. Even though there are many factors affecting the power of a study,  $\alpha$ , effect size and sample size are the most important ones. Even though the calculation of the statistical power is important, it is difficult to say that today's nursing researches commonly determine the number of samples by using power analysis. This study is fairly an introduction knowledge about the power analysis. The article not only conveys some statistical terms regarding the power analysis, but also includes information about how to use the basic information, power analysis types and the power analysis in the sample calculation. Finally, it presents examples from some nursing studies.*

**Keywords:** Power analysis; nursing; research; statistic.

**GİRİŞ**

Bilimsel bilgi elde etme süreci olan araştırma, birbirini izleyen ve etkileyen adım ya da etkinliklerden oluşan sistematik bir süreçtir (Büyüköztürk 2005). Bir araştırma planlarken ilk akla gelen soru, ne kadar denek gerektiğidir (Akgül 2005). Sosyal bilimlerde çoğu araştırma, evrenin tamamı yerine evreni temsil eden, yansız olarak seçilmiş örneklem üzerinden gerçekleşir (Büyüköztürk 2005). İstatistikte kullanılan birçok hipotez testi olduğu (t testi, varyans analizi vb.) ve bu testlerin de her biri için farklı örneklem büyüklüğü formülleri olduğu düşünülürse, örneklem büyüklüğü hesaplamının kolay bir işlem olmadığı söylenebilir. Güç

analizi yapmak, hemşire araştırmacılara örneklem sayısı hakkında bilgi vermenin yanı sıra, çalışmanın bulgularının ne kadar güvenilir olduğu bilgisini de verir.

Cohen, 1962 yılında sosyal psikoloji alanında yapılan araştırmaların, düşük istatistiksel güçte, dolayısı ile zayıf önerilerde bulunduğunu vurguladıktan sonra birçok disiplin, araştırmalarını daha planlanma aşamasındayken, istatistiksel gücü de dikkate alarak düşünmeye başladılar. Hemşire araştırmacılar da çalışmalarında istatistiksel gücün önemini göz önünde bulundurarak hareket ettiler (Coblick 1998). Fakat ülkemizde istatistiksel gücün

\*Atatürk Üniversitesi Sağlık Bilimleri Fakültesi, Hemşirelik Bölümü (Yrd.Doç.Dr.) e-mail: c\_capik36@hotmail.com

yeterince dikkate alındığını söylemek güçtür (Çapık 2013).

Örneklem büyüklüğü belirlenmenin hem ahlaki, hem de istatistiksel yönü vardır (Akgül 2005). Bir çalışma yeterli örneklem sayısına sahip değilse, sonuçların gerçekten istatistiksel olarak önemli olup olmadığını saptamada yetersiz kalınacak, aksi durumda da (örneklem büyüklüğü olması gerekenden yüksek tutulursa) araştırma maliyetinin artması söz konusu olacaktır (Demirel ve Gürler 2010; Kalaycı 2010; Zody 2004). Dolayısı ile gerçekçi bir örneklem sayısı için güç analizi yapılmalıdır. Bu makale güç analizi hakkında giriş niteliğindedir. Makalede, güç analizi ile ilgili bazı istatistiksel terimlerin aktarılmasından sonra, temel bilgilerin, güç analizinin çeşitlerinin, güç analizinin örneklem hesabında nasıl kullanılacağına yönelik bilgilerin verilmesi planlanmıştır. Son olarak da bazı hemşirelik çalışmalarından örnekler verilmiştir.

### GÜÇ ANALİZİ YAPARKEN BİLİNMESİ GEREKEN İSTATİSTİKSEL TERİMLER

İstatistiksel gücü anlayabilmek için, bazı terimleri bilmek gerekir. Sonuçta istatistiksel güç analizi bu terimler üzerinden hesaplanan ve anlam kazanan bir testtir.

#### $H_0$ ve $H_1$ Hipotezleri

Hipotez (varsayım), deneyde henüz yeterli derecede doğrulanmamış ancak doğrulanacağı umulan teorik düşüncedir. Pozitif bilimlerde en çok kullanılan hipotezler sıfır hipotezi ( $H_0$ ) ile alternatif hipotezdir ( $H_1$ ) (Yılmaz 2008).

$H_0$  ve  $H_1$  hipotezlerini örnek üzerinde anlatmak daha kolay anlaşılmasını sağlayabilir. Başlığı "Prostat kanseri hakkında verilen web destekli eğitimin, prostat kanseri bilgi düzeyine etkisi" olan bir çalışmada, olası girişim; bir web sayfası hazırlayarak, katılımcıları bu web sayfası üzerinden eğitmek olacaktır. Bu aşamada, verilen web destekli eğitim, prostat kanseri bilgi düzeyini artırmaz hipotezi; farksızlık,  $H_0$  veya Null hipotezidir. Bir başka anlatımla, girişimin fark yaratmayacağını/etkisiz olacağını savunan hipotezdir (Cozby and Bates 2012; Newman and Kohn 2009).  $H_1$  hipotezi ise bunun tam tersidir, başka bir deyişle, verilen web destekli eğitimin bireylerde prostat kanseri bilgi düzeyini artıracığını savunan hipotezdir (Cozby and Bates 2012). Hemşirelik çalışmalarında, genellikle amaç cümlesinden sonra yer alan araştırmanın hipotezleri bölümünde, hipotezler kurulurken, şu şekilde ifade edilirler:

$H_0$ : Prostat kanserine yönelik verilen web destekli eğitim, bireylerin prostat kanseri hakkındaki bilgi düzeyini artırmaz.

$H_1$ : Prostat kanserine yönelik verilen web destekli eğitim, bireylerin prostat kanseri hakkındaki bilgi düzeyini artırır.

#### Tip I ve Tip II Hata

Hipotezlerin test edilmesinde, iki tip hata söz konusudur. Bu hatalara Tip I ve Tip II hata adı verilir. Tip I Hata ( $\alpha$ ); Gerçekte doğru olan  $H_0$  hipotezinin yanlışlıkla reddedilmesi olasılığıdır (Cozby and Bates 2012; Newman and Kohn 2009). Önceki örneğimiz üzerinden gidilecek olursa,  $H_0$  hipotezi (Prostat kanserine yönelik verilen web destekli eğitim, bireylerin prostat kanseri hakkındaki bilgi düzeyini artırmaz) gerçekte doğru iken, yanlış bir şekilde bu hipotezi reddetme ihtimalidir. Yani yanlış olarak "bilgi düzeyi artmıştır" sonucuna varılması ihtimalidir (Tablo 1).

**Tablo 1.** Tip I ve Tip II Hata Kavramları\*

ARAŞTIRMA SONUCU	GERÇEK DURUM	
	Web destekli eğitim bilgi düzeyini artırır.	Web destekli eğitim bilgi düzeyini artırmaz.
Web destekli eğitim bilgi düzeyini artırır ( $H_1$ ).	<b>Doğru karar</b>	Tip I Hata ( $\alpha$ )
Web destekli eğitim bilgi düzeyini artırmaz ( $H_0$ ).	Tip II Hata ( $\beta$ )	<b>Doğru karar (Güç)</b> ( $1 - \beta$ )

\* Akgül 2005'den uyarlanmıştır.

Tip I hata araştırmacı tarafından önceden belirlenir ve genellikle 0.05 ya da 0.01 olarak seçilir (Cozby and Bates 2012). Eğer Tip I hata düzeyi 0.05 olarak kabul edilirse, fark yoktur sonucu çıktığında %95 olasılıkla fark yoktur anlamına gelir.

Tip II Hata ( $\beta$ ): Gerçekte doğru olan  $H_1$  hipotezini kabul etmeyerek, reddedilmesi durumunda ortaya çıkan hatadır.  $\beta$  ile gösterilir (Newman and Kohn 2009). Önceki örneğe dönülecek olursa,  $H_1$  Hipotezi (Prostat kanserine yönelik verilen web destekli eğitim, bireylerin prostat kanseri hakkındaki bilgi düzeyini artırır) gerçekte doğru iken ve kabul edilmesi gerekirken, hatalı olarak "eğitim bilgi düzeyini artırmamıştır" sonucuna varılması durumudur (Tablo 1). Genel olarak, Tip II hata düzeyi 0.20 olarak belirlenir. Diğer bir ifade ile gruplar

arasında olan farkı saptayamama ihtimali %20 olarak kabul edilir (Akgül 2005). İleride anlatılacağı üzere,  $1 - \beta$  Çalışmanın gücünü verir (Newman and Kohn 2009). Tip II hatanın 0.20 olarak kabul edildiği bir durumda çalışmanın gücü (1-0.20) %80 olur (Akgül 2005). Tip II hata ve güç arasında negatif bir ilişki vardır (Goodwin 2010).

### p Değerinin Anlamı

p değerini hesaplamak için  $X^2$  veya t testi gibi hipotez testleri kullanılır. Hangi testin seçileceğine ise, verinin türüne göre karar verilir. Klasik olarak eğer p değeri  $\alpha$  değerinden yani Tip I hatadan küçükse sonuç önemli kabul edilir ve  $H_0$  hipotezi reddedilir. Yaygın olarak araştırma bulguları verilirken,  $H_0$ ,  $H_1$  veya Tip I hata terimleri kullanılmaz ve p değeri 0.05'den küçükse sonuç/gruplar arasındaki fark anlamlıdır şeklinde rapor edilir (Cozby and Bates 2012; Newman and Kohn 2009). Kısacası, eğer  $p = 0.05$  ise, %5 olasılıkla  $H_0$  hipotezi yanlışlıkla reddedilmiştir ve %95 olasılıkla fark gerçektir anlamına gelir. Başka bir deyişle,  $H_0$  hipotezinin % türünden doğru olma ihtimalidir (Newman and Kohn 2009).

### P Değeri Neden Anlamsız Çıkar?

Araştırmacılar genellikle  $H_0$  hipotezi ile ilgilenmezler ve önemlilik tespit etmeyi isterler. p değerinin anlamsız çıkmasının birkaç nedeni olabilir. Tabii ki ilk neden sonuçların gerçekten anlamsız olmasıdır. Bunun yanı sıra, bazen gerçekte önemlilik olsa bile araştırmacılar fark yoktur sonucuna ulaşırlar, bu durum zaten Tip II hatanın yapılması durumudur. Bazen Tip II hatanın nedeni, uygulanan girişim prosedüründen kaynaklanır. Bazen de örneklem büyüklüğü ve etki büyüklüğünden kaynaklanabilir. Örneklem büyüklüğü gerçek etkiyi ortaya çıkaracak yeterlilikte değildir veya çok küçük etki büyüklüğünü tespit edecek genişlikte olmayabilir (Cozby and Bates 2012).

### Etki Büyüklüğü

Etki büyüklüğü, son yıllarda eğitim alanında yapılan araştırmalarda önem verilen bir kavram olarak karşımıza çıkmaktadır. Aynı zamanda APA tarafından, araştırmalarda p değeri ile birlikte mutlaka etki büyüklüğünün de hesaplanarak raporlanması gerektiği belirtilmektedir (Beins and McCarthy 2012; Özsoy ve Özsoy 2013; Jackson 2011). Özsoy ve Özsoy'un (2013) eğitim bilimlerinde yaptıkları bir araştırmaya göre Türkiye'de yapılan çalışmalarda etki büyüklüğünün rapor edilme oranı %7.2 olarak saptanmıştır.

Etki büyüklüğü, ortalamalar arasındaki farkın standardize edilmiş bir ölçümüdür. Genelde standardize fark olarak rapor edilir. Eğer bir korelasyon analizi ise standardize edilmiş müşterek değişkenlik olarak da bahsedilebilir (Beins and McCarthy 2012). Tıp literatürde, yapılan bir girişimin etkisinin büyüklüğünün ölçümüdür (Beck 2006). Bir çalışmada normal hipotez testleri iki grup arasında fark saptayabilir, fakat sonuçlar araştırmacıya bu farkın boyutu/büyüklüğü hakkında bilgi vermez. Etki büyüklüğü standardize edilmiş farkın bir ölçütü olduğu için bu bilgiyi verebilir (Goodwin 2010).

Bir örnekle açıklanacak olursa, iki guruptan birisinin bilgi ortalaması  $12.01 \pm 3.20$ , diğersinin  $15.20 \pm 1.50$  ise ve arada istatistiksel olarak anlamlı fark var ise yaklaşık 3 puanlık bilgi farkı gerçekte dikkate değer bir fark mıdır? Yoksa uygulamada bir anlam ifade etmiyor mu? Sorusunun cevabıdır.

Farklı çalışma dizaynlarında etki büyüklüğünü hesaplama yolları da farklıdır (Goodwin 2010). Her test için başka formülle hesaplanır. Örneğin t testinde iki grup ortalaması karşılaştırılır, bu iki grup arasındaki standardize edilmiş farkın büyüklüğünün hesaplanması için gerekli formül:

$$d = \frac{M_1 - M_2}{Sp}$$

Şeklinde (Beins and McCarthy 2012; Özsoy ve Özsoy 2013). Burada,

$d$  : Cohen's d sini yani bu durumda etki büyüklüğü değerini,

$M_1$  : Deney grubu ortalaması

$M_2$  : Kontrol grubu ortalamasını

$Sp$  : Harmanlanmış (pooled) standart sapmayı temsil eder.  $Sp$ 'nin formülü ise aşağıdaki gibidir.

$$Sp = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2}}$$

$n$  : Birey sayısı,

$S_1^2$  : Birinci grubun standart sapması,

$S_2^2$  : İkinci grubun standart sapmasıdır.

Etki büyüklüğünü t testi için hesaplamak kolaydır. Bir grubun ortalaması  $50 \pm 5.60$  ( $n=10$ ), diğer grubunki  $30 \pm 3.74$  ( $n=10$ ) ise, yukarıdaki formülde sayılar yerine konulduğunda, etki büyüklüğü 4.52 çıkacaktır. Tablo 2'de görüldüğü gibi t testinde standardize edilmiş küçük etki büyüklüğü 0.2, orta etki büyüklüğü 0.5 ve büyük

etki büyüklüğü 0.8'dir (Cohen 1992). Bizim bulduğumuz değer (4.52) Tablo 2'ye göre yorumlandığında oldukça büyük bir etki büyüklüğü olduğu görülür. Diğer hipotez testleri için (örneğin varyans analizi) formül değişiklik gösterir ve karmaşıklaşabilir, bundan dolayı etki büyüklüğü birçok istatistik programı vasıtası ile hesaplanabilir.

**Tablo 2.** Cohen Tarafından Bildirilen Etki Büyüklüğü Aralıkları\*

	Düşük Etki Büyüklüğü	Orta Etki Büyüklüğü	Yüksek Etki Büyüklüğü
Student t-testi	0.2	0.5	0.8
Varyans analizi	0.1	0.25	0.4
Ki-kare testi	0.1	0.3	0.5

\* Cohen 1992

Etki büyüklüğü 3 yolla önceden tahmin edilebilir. Birincisi benzer çalışmalar daha önceden yapılmış ise o çalışmanın standardize etkisini örnek olarak yapılabilir. İkincisi, tümden gelimli bir yaklaşımla literatürde yapılacak girişimin etkisi belli ise ona benzer bir etkiyi beklemek, üçüncüsü ise Tablo 2'de görülen önceden standardize edilmiş, düşük, orta ve yüksek etki büyüklüklerini kullanmaktır (Murphy and Myers 2004).

### Örneklem Büyüklüğü

Bu bölümde örneklem büyüklüğünün nasıl hesaplanacağından çok, örneklem büyüklüğü ve istatistiksel güç arasındaki ilişki ifade edilmeye çalışılmıştır. Örneklem büyüklüğü gücü etkileyen önemli bir faktördür ve her bir örnek üzerinde yapılan gözlemlerin sayısı olarak tanımlanabilir. Deneysel çalışmalarda, üzerinde işlem uygulanan örnek sayısı örneklem büyüklüğünü ifade eder. Örneklem grupları eşit sayıda gözlem içeriyorsa elde edilen veriler dengelidir. Aksi halde grupların örneklem büyüklükleri dengesiz olacaktır. Örneklem büyüklüğünün artırılması gücün artmasını sağlar (Yılmaz 2008).

Bir araştırma planlanırken ilk sorulan soru, örneklem büyüklüğünün nasıl belirleneceği ve ne kadar olacaktır (Akgül 2005). Bu soruya kolayca cevap vermek her zaman mümkün olmamaktadır. Henüz planlama aşamasında iken, güç analizini kullanmak çalışma için kaç katılımcı gerektiği hakkında hassas karar vermeye yardımcı olur (Murphy and Myers 2004). Güç analizi yaparak gerekli örneklem

sayısı hakkında bilgi edinmek pratik bir yol olmasına rağmen, hemşirelik çalışmalarında istatistiksel güç yaygın olarak hesaplanmamaktadır (Çapık 2013). Çalışmaya gerçek durumu ortaya çıkaracak kadar yeterli genişlikte örneklem alınmadığında Tip II hata ( $\beta$ ) büyür, dolayısı ile güç ( $1-\beta$ ) küçülür. Örneğin bir hastalığın görünme ihtimali %10 ise biz 50 kişi ile çalışırsak bu 50 kişi içinde hastalık pozitif kişi yakalanmayabilir ve sonuç negatif çıkar, dolayısı ile gerçekte var olan bir şey, yoktur şeklinde yorumlanır, bu durumda hem yanlış bir sonuç elde etmiş oluruz, hem de güç küçülmüş olur (Beins and McCarthy 2012). 500 kişiye göre 5000 kişi ile yapılan bir çalışmanın yüksek güçte çıkma olasılığı çok daha fazladır. Fakat bu 5000 kişi üzerinde yapılan çalışmada 0.03 mmHg kan basıncı düzeyi bile anlamlı çıkabilir (Shoben 2009). Bu küçük farkın klinik olarak bir anlamı olmaması (etki büyüklüğünün düşük olması) ise başka bir konudur.

### İSTATİSTİKSEL GÜÇ

Sosyal ve davranış bilimlerinde istatistiksel analizler iki amaca hizmet eder, birincisi ilgili çalışmadan neler elde edildiğini tanımlayıcı olarak göstermek, ikincisi istatistiksel testler yaparak sonuçlardan bir çıkarım yapmaktır. Fakat buradaki sıkıntı mevcut örneklem üzerinden elde edilen bulguların tüm popülasyon için uygun olup olmadığını bilmemektir. Çalışmalarda düşük güç genellikle yanlış yorumlamalara sebep olur. Böyle araştırmaların sonuçları da genel popülasyon için yanlış bilgileri içerebilir. (Murphy and Myers 2004)

Güç analizi veri toplamaya başlanmadan önce yapılabildiği gibi veri toplandıktan sonra da yapılabilir. Deney öncesinde yapılan güç analizi, hedeflenen gücü elde etmek amacıyla gerekli olan örneklem büyüklüğünü tahmin etmede kullanılır. Veri toplama sonrasında yapılan güç analiziyle de çalışmanın gerçek gücü hesaplanabilir.

İstatistiksel güç,  $H_0$  hipotezini doğru olarak reddetme ihtimalidir (Beins and McCarthy 2012; Dorey 2011). İstatistiksel güç, popülasyonda var olan farklılığı bulabilme olasılığıdır. Tip II hatanın, gerçekte yanlış olan  $H_1$  hipotezini reddetmeyerek, kabul etmek olduğu hatırlandığında (Newman and Kohn 2009), **istatistiksel güç gruplar arasında anlamlı bir fark saptanması durumunda bu farkın gerçek olabilme ihtimalidir** (Goodwin 2010). Bir örnekle açıklamak gerekirse, iki grup ortalaması arasındaki fark anlamlı olarak

saptanmışsa, yani  $H_1$  hipotezi kabul edilerek “gruplar arasında fark vardır” sonucuna varılmışsa, bu sonucun doğru olma ihtimali o testin gücüdür (Tablo 1’den takip edilebilir). Eğer araştırmacı gruplar arasındaki farkı ortaya çıkaracak yeterli sayıda örnekleme çalışmamışsa, gruplar arasında gerçekte var olan farkı saptayamayacaktır ve test gücü düşük olacaktır.

Güç tıpkı p değeri gibi şartlı bir olasılıktır (Dorey 2011). Yani başka etkenlere göre değişir. İstatistiksel gücün nelerden etkilendiği ise önemli bir konudur. Güç büyük oranda  $\alpha$ , etki büyüklüğü ve örneklem büyüklüğüne bağlı olarak değişir (Cohen 1992). Eğer bunlardan herhangi ikisi biliniyorsa diğer üçüncüsü hesaplanabilir (Dorey 2011). Gücü etkileyen etmenlerden ilki olan  $\alpha$  ne kadar küçük olursa,  $\beta$  o kadar büyür, böylece, testin gücü,  $1-\beta$  küçülür (Akgül 2005). Genel olarak p değeri 0.05’den daha küçük bir değer olarak sabitlenirse güç daha düşük çıkar (Shoben 2009). Gücü etkileyen diğer bir öge olan etki büyüklüğünün yüksek olması ortalamalar arasındaki standardize edilmiş farkın çok olduğu anlamına gelir, bu durumda farkı tespit etmek kolaylaşır ve güç büyür (Beins and McCarthy 2012). Gücü etkileyen son önemli öge ise örneklem sayısıdır. Etki büyüklüğü küçük çıkarsa bunu değiştirmek zordur. Fakat gücün düşük çıkması durumunda örneklem sayısı yükseltilebilir. Zaten güç, örneklem sayısından etkilendiği için araştırmacılar en uygun örneklem sayısını güç analizi ile hesaplayabilmektedirler. Örneklem sayısı yükseltilebilir fakat burada göz ardı edilen bir konu vardır, oda örneklem sayısı çok aşırı yükseltirse, ortalamalar arasındaki çok küçük farklılıkların bile istatistiksel olarak anlamlı çıkma eğiliminde olmalarıdır. Örneğin her biri 5000 kişiden oluşan 2 gruba portakal suyunun IQ düzeyleri üzerine etkisi karşılaştırıldığında, arada fark saptanırsa ve IQ düzeyi sadece 2 puanlık bir farktan dolayı anlamlı çıkarsa, bu iki puan gerçekte hiçbir anlam ifade etmeyebilir (Goodwin 2010). Yani şöyle bir çıkarım yapılabilir. Örneklem sayısı az olduğu için güç düşük çıkabilir, bu durumda örneklem hacmi artırılarak güç yükseltilebilir. Böyle bir durumda da gruplar arasındaki çok küçük farklar istatistiksel olarak önemli, etki büyüklüğü de düşük olabilir. Sonuçta elde ettiğimiz bulgular gerçek hayatta bir işe yarayabilir.

Hipotez testlerinde  $H_0$  yani farksızlık hipotezi kabul edilirse gruplar arasında fark

yoktur sonucuna ulaşılır ve normal olarak tartışma bu yönde yapılır, ama bu çalışmanın gücü düşükse aslında gerçekte var olan farkı biz görmemiş oluruz ve yanlış bulgu ve yorum yazmış oluruz (Murphy and Myers 2004). Bu durum örneklem büyüklüğü hesaplamının yanı sıra güç analizini önemli yapan başka bir etkidir. Bu durumu gösterir bir örnek, Tablo 3’de sunulmuştur.

**Tablo 3.** Deneysel ve Kontrol Gruplarında Ağrı Puanlarının Karşılaştırılması

Grup	n	Ort.	SS	Önemlilik
Kontrol	30	10.00	5.12	t=0.317
Deneysel	31	9.68	2.68	p=0.752

Yukarıdaki tabloya güç analizi yapmadan, klasik hipotez testinin (t testi) p değerine göre karar verirse, iki grup arasında fark olmadığı çok net olarak görülmektedir. Fakat bu testin gücü hesaplandığında gücün %6 olduğu görülür. Dolayısı ile arada fark olmadığına inanmak büyük ihtimalle bizi yanıltmıştır. Daha büyük örnekleme başka sonuçlar çıkaracaktır.

Araştırmacılar, çalışmaya başlamadan önceki güç analizini (priori) örneklem büyüklüğü,  $1-\beta$  ve etki büyüklüğünü kestirmek için kullanırlar (Taylor 2008). Bu durumda Tablo 2’de gösterilen standart etki büyüklükleri kullanılabilir (Cohen 1992). Bunun yanı sıra tamamlanmış bir çalışma için de sonradan güç analizi yapılabilir. Çalışmadan sonraki güç analizi (post hoc) ise mevcut çalışmada gözlenen  $1-\beta$  ve etki büyüklüğünü öğrenmek için kullanılır. Post hoc güç analizi; anlamsız sonuçlar bulunduğu, literatürle çelişen bulgular edinildiğinde veya bulunan farkın klinikteki anlamlılığı tartışıldığında daha sık yapılır (Taylor 2008).

İstatistiksel güç hesabı yapıldığı zaman raporu genellikle materyal ve metod bölümünde, örneklem genişliği başlığı altında yer alır (Newman and Kohn 2009).

### **Bir Çalışmanın İstatistiksel Gücü Kaç Olmalıdır?**

Güç 0.00 ve 1.00 arasında değişir ve güç yükseldikçe ölçümün farkı bulmak için daha hassas olduğu anlamına gelir. Gücün 0.80 ve üzerinde olması istenir (Pagano 2009). Bazı yazarlar ise bu sınırı 0.70 - 0.90 arasında bildirmektedir (Cozby and Bates 2012). Bir çalışmanın gücü 0.50’den düşükse sonuçları

genellikle yanlış yorumlanır (Murphy and Myors 2004).

Yukarıda açıklandığı gibi, yüksek etki büyüklüğü ve geniş örneklem grubu istatistiksel gücün fazla çıkmasını sağlar (Dorey 2011). Bunların yanı sıra bazı diğer faktörler de çalışmanın gücünü etkileyebilir. Örneğin; gruplarda eşit sayıda denek bulunması, çalışmaya alınan toplam denek açısından en büyük güç değerinin elde edilmesini çoğunlukla sağlar (Akgül 2005; Taylor 2008). Bunun yanı sıra bağımlı değişkenlerin analizinde pre ve post testler genelde aynı kişilerdir ve gücün yüksek çıkmasına olumlu katkı yapar (Beins and McCarthy 2012; Taylor 2008). Popülasyonun normal dağılım sergilemesi deneyin gücü üzerinde etkilidir. Normal dağılımda meydana gelecek sapmalar gücü düşürmektedir. İstatistiksel analizde kullanılan yöntemler de güc üzerine etkili olabilir. Tek yönlü testler çift yönlü testlere oranla genellikle daha güçlüdür (Taylor 2008; Yılmaz 2008).

#### **Güç Analizinin Hemşirelik Çalışmalarında Kullanımı**

Yapılan literatür taramasında güncel olan ve hemşirelik çalışmalarında güç analizi yapmış araştırmaları inceleyerek rapor eden yalnızca iki makaleye ulaşılmıştır. Ulaşılan makalelere bakılarak karar vermek zor olsa bile hemşirelik çalışmalarında güç analizinin çok kullanılmadığını belirtmek gerekir. Bu iki çalışmadan ilki Gaskin ve Happell (2013) tarafından yapılmıştır. Yazarlar bir hemşirelik dergisinde yayınlanan makalelerin sadece %17'sinde istatistiksel güce ilişkin bilgiler bulunduğunu belirlemişlerdir. Yazarlar ayrıca orta etki büyüklüğünde makalelerin %79'unun istenilen güce ulaşabildiğini belirtmiştir (Gaskin and Happell 2013). Diğer çalışma ise Çapık (2013) tarafından yapılmıştır. Çapık (2013) Türkiye'de bir hemşirelik dergisinde son 3 yılda yayınlanan 61 araştırma makalesini incelemiş ve sadece bir çalışmada güç analizi yapılarak metin içinde verildiğini belirlemiştir. Çapık'ın (2013) çalışmasına göre parametrik testlerin kullanıldığı

durumlarda istatistiksel güçler daha yüksektir ve orta etki büyüklüğünde testlerin sadece %46.4'ünün, istatistiksel gücünün 0.80 ve üzerinde olduğunu belirlemiştir.

#### **Güç Analizi Yapabilen Yazılımlar**

Birçok bilgisayar programı etki büyüklüğü, önemlilik düzeyi ve bazı bilgiler girildiğinde istatistiksel gücü hesaplayabilmektedir (Cozby and Bates 2012). Bunlardan bazıları şunlardır; PASS, Gpower, nQuery, UnifyPow, SPSS (modülü yüklendiğinde), Power and Precision ve STATISTICA.

Bu programların yanı sıra birçok çevrimiçi çalışan ve güç analizi yapabilen WEB sitesine internetten ulaşılabilir.

#### **SONUÇ**

İstatistiksel güç, gruplar arasında anlamlı bir fark saptanması durumunda bu farkın gerçek olabilme ihtimalidir (Goodwin 2010). Güç analizi hemşire araştırmacılara örneklem sayısı hakkında bilgi vermenin yanı sıra, çalışmanın bulgularının ne kadar güvenilir olduğu bilgisini de verir. Güç analizi bir çalışmaya başlamadan önce yapılabildiği gibi, çalışma tamamlandıktan sonrada yapılabilir. Güç büyük oranda  $\alpha$ , etki büyüklüğü ve örneklem büyüklüğüne bağlı olarak değişir (Cohen 1992). Güç, örneklem sayısından etkilendiği için araştırmacılar en uygun örneklem sayısını güç analizi ile hesaplayabilirler. Güç 0.00 ve 1.00 arasında değişir ve güç yükseldikçe ölçümün farkı bulmak için daha hassas olduğu anlamına gelir. Gücün 0.80 ve üzerinde olması istenir (Pagano 2009). Ulaşılan makalelere bakılarak karar vermek zor olsa bile hemşirelik çalışmalarında güç analizinin çok kullanılmadığını belirtmek gerekir.

Bir çalışma tamamlandıktan sonra, çalışmanın sınırlılıklarını düzeltmek zordur. Az örneklem sayısı nedeniyle, yanıltıcı sonuçların verilmesi ise faydadan çok zarar getirebilir. Bu nedenle en azından az örneklem sayısına bağlı yanıltıcı sonuçların verilmesine güç analizi yaparak engel olunabilir.

#### **KAYNAKLAR**

**Akgül A.** Tıbbi Araştırmalarda İstatistiksel Analiz Teknikleri SPSS Uygulamaları. Ankara: Emek Ofset; 2005. p.41-55.

**Beck CT.** Meta-Analysis. In: Fitzpatrick JJ, Wallace M, editors. Encyclopedia of Nursing Research. 2 ed. New York: Springer Publishing Company; 2006. p. 340.

**Beins BC, McCarthy MA.** Research Methods and Statistics. New Jersey: Pearson Education, Inc.; 2012. p.149-71.

**Büyükoztürk Ş.** Sosyal Bilimler için Veri Analizi El Kitabı. Ankara: Pegem A Yayıncılık; 2005. p.65-9.

**Coblick GE.** Statistical Power Analysis of Nursing Research. Alabama: Auburn University Press; 1998. p.52-78.

- Cohen J.** Statistical Power Analysis. *Current Directions in Psychological Science* 1992;1(3):98-101.
- Cozby P, Bates S.** *Methods in Behavioral Research*. New York: McGraw-Hill; 2012. p.262-74.
- Çapık C.** Bir Hemşirelik Dergisinde Yayınlanan Makalelerde İstatistiksel Güçlerin İncelenmesi. *Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi* 2013;16(3):170-5.
- Demirel, N, Gürler, S.** Practical Approaches to the Determination of the Sample Size in Clinical Studies. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi* 2010;16(2):205-11.
- Dorey FJ.** Statistics in Brief: Statistical Power: What is it and when should it be used? *Clin Orthop Relat Res* 2011;469(2):619-20.
- Gaskin CJ, Happell B.** Power of Mental Health Nursing Research: A Statistical Analysis of Studies in the *International Journal of Mental Health Nursing*. *International Journal of Mental Health Nursing* 2013;22(1):69-75.
- Goodwin CJ.** *Research in Psychology Methods and Design*. Sixth ed. New Jersey: John Wiley & Sons, Inc; 2010. p.157-61.
- Jackson SL.** *Research Methods: A Modular Approach*. Second Edition ed. Belmont: Linda Schreiber-Ganster; 2011. p.187-91.
- Kalaycı, S.** SPSS Uygulamalı çok değişkenli istatistik teknikleri. Ankara: Asil Yayınevi; 2010. p. 221–31.
- Murphy KR, Myers B.** *Statistical Power Analysis, a Simple and General Model for Traditional and Modern Hypothesis Test*. London: Lawrence Erlbaum Associates; 2004. p.55-68.
- Newman T, Kohn M.** *Evidence-Based Diagnosis*. New York: Cambridge University Press; 2009. p.220-39.
- Özsoy S, Özsoy G.** Eğitim Araştırmalarında Etki Büyüklüğü Raporlanması. *İlköğretim Online* 2013;12(2):334-46.
- Pagano RR.** *Understanding Statistics in the Behavioral Sciences*. Belmont: Wadsworth, Cengage Learning; 2009. p.267-77.
- Shoben A.** Interpreting Hypothesis Tests. In: Adeney KL, Weiss NS, Shoben AB, editors. *Epidemiology and Biostatistics An Introduction to Clinical Research*. New York: Springer Science-Business Media, LLC; 2009. p. 181-90.
- Taylor SE.** Hypothesis Testing. In: Boslaugh S, McNutt L-A, editors. *Encyclopedia of epidemiology*. California: SAGE Publications, Inc.; 2008. p. 515-16.
- Yılmaz Y.** Deneylerde İstatistiksel Güç ve Gücü, Etkileyen Faktörler. *Akademik Gıda* 2008;6(6):45-50.
- Zodpey, SP.** Sample Size and Power Analysis in Medical Research. *Indian J Dermatol Venereol Lepro* 2004; 170(2):123-8.