

ARAŞTIRMA

BİR HEMŞİRELİK DERGİSİNDE YAYINLANAN MAKALELERDE İSTATİSTİKSEL GÜÇLERİN İNCELENMESİ*

CANTÜRK ÇAPIK**

Alınış Tarihi: 13.03.2012

Kabul Tarihi: 28.05.2013

ÖZET

Amaç: Bir araştırmaya başlarken, yeterli örneklem sayısını hesaplamak gerekir. Örneklem sayısını hesaplamanın yollarından birisi de güç analizi yapmaktır. Bu çalışmanın amacı bir hemşirelik dergisinde 2009 – 2011 yılları arasında yayınlanmış olan makalelerde kullanılan istatistiksel karşılaştırmaların güç analizini yapmak ve sonuçlarını sunmaktır.

Yöntem: Tanımlayıcı türde yapılan bu çalışmada 61 araştırma makalesi incelenmiş, bu makalelerde kullanılan 725 istatistiksel testin güç analizi yapılmıştır. Güç analizleri Cohen tarafından bildirilen düşük, orta ve yüksek etki büyüklüğüne göre incelenmiş ve her test için üç farklı sonuç verilmiştir. Güç analizi yapabilmek için incelenen makalelerde kullanılan istatistiksel testleri, ortalamaları, standart sapmaları, çalışmaya alınan örneklem sayılarını ve diğer bilgileri not edebilmek için bir “yayın tanıtım formu” kullanılmıştır. Bu form araştırmacı tarafından oluşturulmuştur. İstatistiksel çözümlemede SPSS 17.0 paket programları kullanılmıştır.

Bulgular: İncelenen makaleler içinde sadece bir araştırmanın istatistiksel gücü hesaplanarak metin içinde verilmiştir. Makalelerde kullanılan parametrik testlerin güçleri, non-parametrik testlerin güçlerinden daha yüksektir. Orta etki büyüklüğünde testlerin %46.4’ünün, yüksek etki büyüklüğünde ise %88.7’sinin istatistiksel gücü 0.80 ve üzerindedir.

Sonuç: Çalışmanın sonuçlarına göre hemşirelik araştırmalarında güç analizi yaparak örneklem sayısını belirlemek yaygın değildir. İncelenen çalışmaların bir kısmının istenilen güce ulaşması için daha fazla örnekleme ihtiyacı vardır. Bu nedenle hemşire akademisyenlere araştırmalarına başlamadan önce istatistiksel güç hesaplayarak örneklem büyüklüğünü belirlemeleri önerilmektedir.

Anahtar Kelimeler: Hemşirelik; güç analizi; örneklem sayısı.

ABSTRACT

Examination of Statistical Powers in Articles Published in A Nursing Journal

Purpose: It is required to calculate sufficient number of samples while commencing a study. One of ways of calculating number of samples is to perform power analysis. The purpose of this study is to conduct the power analysis of statistical comparisons used in articles published in a nursing journal between 2009-2011 and present the results descriptively.

Method: This descriptive study examined 61 research articles. Power analysis of 725 statistical tests used in these articles was performed. Power analyses were examined by Cohen according to low, moderate and high effect size and three different results were given for each test. In order to conduct the power analysis, a “publication description form”, was used to note statistical tests used in the articles, means, standard deviations, number of samples and other information. SPSS 17.0 software programs were used to perform the statistical analysis.

Results: Among the articles examined, the statistical power of only one research was calculated and presented in the text. Powers of parametric tests used in articles are higher compared to non-parametric tests. The statistical powers of 46.4% of tests at moderate-effect size and 88.7% of tests at high-effect size are 0.80 and above.

Conclusion: According to results of the study, it is not common in nursing studies to determine the number of samples by conducting power analysis. A greater number of samples is required for some of studies examined to reach the intended power. Thus, it is suggested for nurse academicians to determine the sample size by calculating the statistical power before the study.

Keywords: Nursing; power analysis; number of samples.

*Bu çalışma, “The 1st International Clinical Nursing Research Congress” (29 Mayıs-1 Haziran 2012, Kaya Hotel&Convention, İzmir, Türkiye)’de poster bildiri olarak sunulmuştur.

**Kafkas Üniversitesi Kars Sağlık Yüksekokulu, Hemşirelik Bölümü (Yrd.Doç.Dr) c_capik36@hotmail.com

GİRİŞ

Örneklem büyüklüğünün belirlenmesi ve istatistiksel güç hesabı birçok nedenden dolayı önemlidir. Örneğin birçok akademik dergi çalışmanın gücünü istemektedir ve etik komiteler örneklem sayısı hakkında açıklama isteyebilirler (Parajapati, Dunne and Armstrong 2010). 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; Zodpey 2004). Bazı klinik durumlarda ise çok sayıda deneye ulaşmak gerçekçi olmaktan uzak olabilir (Usami 2011).

Güç analizi yapmak, örneklem sayısını hesaplamamanın yollarından birisidir, örneklem sayısının kaç olması gerektiği hakkında bilgi vermekle birlikte, tamamlanmış bir çalışmanın yeterli örneklem sayısına sahip olup olmadığını da belirlemeyi sağlar. İstatistiksel güç, yanlış bir sıfır hipotezini doğru olarak reddetme ihtimalidir (Gravetter and Wallnau 2009). Bir örnekle açıklamak gerekirse, iki grup ortalaması arasındaki fark anlamsız olarak saptanmışsa, yani sıfır hipotezi kabul edilerek “gruplar arasında fark yoktur” sonucuna varılmışsa, bu sonucun doğru olma ihtimali o testin gücüdür. Eğer araştırmacı gruplar arasındaki farkı ortaya çıkaracak yeterli sayıda örnekleme çalışmamışsa, gruplar arasında gerçekte var olan farkı saptayamayacaktır ve test gücü düşük olacaktır.

Güç analizini daha iyi ifade edebilmek için hipotez testlerinde yapılabilecek iki tip hatanın bilinmesi gerekir. Bunlardan birincisi, araştırmacının gerçekte kabul edilmesi gereken sıfır (farksızlık) hipotezini yanlış bir şekilde reddetmesi (Tip I hata), ikincisi ise, gerçekte reddedilmesi gereken sıfır hipotezinin yanlış olarak kabul etmesidir (Tip II hata) (Coolican 2013). Tip I hata “ α ” ile tip II hata ise “ β ” ile gösterilir. Gruplar arasındaki farkı saptamaya çalışan istatistiksel testler, gruplar arasında “fark yoktur” u savunan sıfır hipotezini reddetmeye çalışır. Eğer sıfır hipotezi reddedilirse bu gruplar arasında fark olduğu anlamına gelir (Healey 2009). Bir testin istatistiksel gücü ise, $1-\beta$ 'yi temsil eder (Coolican 2013).

Bir çalışmanın istatistiksel gücü, önemlilik düzeyi, etki büyüklüğü ve örneklem sayısından etkilenir (Gravetter and Wallnau 2009). Bu üç parametre biliniyorsa dördüncü

parametre olarak ta güç hesaplanabilir (Dorey 2011). İstatistiksel önemlilik yani “p” değeri tip I hata yapma olasılığını verir. Anlamlılık seviyesi diğer alanlarda olduğu gibi hemşirelik araştırmalarında genellikle 0.05 olarak alınır (Coolican 2013). Bu değer güven aralığının %95 olduğunu yani, test edilen değer %95 güven aralığında kalırsa anlamlılığın kabul edileceği anlamına gelir (Kalaycı 2010). Önemlilik düzeyi düştükçe gücü yüksek tutmak için örneklemin artırılması gerekir (Parajapati, Dunne and Armstrong 2010). Etki büyüklüğü ise, en kısa tanımla iki ortalama arasındaki farkın büyüklüğünü ifade etmektedir. Bu fark kullanılan testin türüne göre bir formülle hesaplanır (Lan and Lian 2010). Etki büyüklüğü arttıkça tip II hata azalır. Etki büyüklüğünü belirlemek için bir pilot çalışma yapılabilir, uzman birisinin tahminlerine güvenilebilir veya Cohen tarafından belirlenmiş standart büyüklükler kullanılabilir (Parajapati, Dunne and Armstrong 2010). Cohen (1988), çeşitli örneklem ve veri türlerinde çalışarak, etki büyüklüğü aralıklarını sınıflandırmıştır. Bu sınıflandırmaya göre, etki büyüklüklerini düşük, orta ve yüksek olarak isimlendirmiştir. Çeşitli istatistiksel testlerde etki büyüklükleri Tablo 1’de sunulmuştur.

Tablo 1. Cohen Tarafından Bildirilen Etki Büyüklüğü Aralıkları

Test	Düşük Etki Büyüklüğü	Orta Etki Büyüklüğü	Yüksek Etki Büyüklüğü
Bağımsız gruplarda t-testi	0.2	0.5	0.8
Varyans Analizi	0.1	0.25	0.4
X ² Testi	0.1	0.3	0.5

*(Cohen 1988)

Etki büyüklüğünün önemi çalışmadan çalışmaya değişebilir. Örneğin, bir tedavi yöntemi ölümleri %1 oranında azaltıyorsa bu klinik olarak önemlidir, fakat bir tedavi yönteminin geçici astımı %20 oranında azaltması klinik olarak göz ardı edilebilir (Zodpey 2004). Örneklem büyüklüğü ise, istatistiksel gücü hesaplayabilmek için gereken diğer bir kavramdır. Örneklem büyüklüğü, çalışmaya alınan toplam birey sayısıdır ve örneklem büyüklüğü arttıkça testin gücünde artar. Güç, etki büyüklüğünün ve örneklem hacminin bir fonksiyonu olduğu için, düşük etki büyüklüğünde gücü yüksek tutmak için örneklem sayısının fazla olması gerekir (Gravetter and

Wallnau 2009). Çalışmalarda istatistiksel gücün 0.80 olmasının yeterli olduğunu bildirmektedir. Bu ortalamalar arasında fark saptanmamışsa, bu sonucun %80 ihtimalle doğru olduğu anlamına gelir (Cohen 1988; Parajapati, Dunne and Armstrong 2010).

Evrenin bilindiği veya bilinmediği durumlarda gerekli örneklem sayısını belirlemek için çeşitli formüller bulunmaktadır (Healey 2009). Fakat bu formüllerde yerine koymak için gerekli verilere ulaşmak zaman alıcı veya imkânsız olabilir. Güç analizi yapmak hemşire akademisyenlere bir çalışmaya başlamadan önce kaç kişilik bir grupta çalışmalarını gerektiği hakkında daha kolay bilgi verebilir. Bu durumda standart etki büyüklüklerini kullanmak gerekir (Cohen 1988). Bunun yanı sıra tamamlanmış bir çalışma için de sonradan güç analizi yapılabilir (Wang 2010). Güç analizi yaparak gerekli örneklem sayısı hakkında bilgi edinmek pratik bir yol olmasına rağmen, hemşirelik çalışmalarında istatistiksel güç yaygın olarak hesaplanmamaktadır. Gaskin ve Happell (2013) bir hemşirelik dergisinde yayınlanan makalelerin sadece %17'sinde istatistiksel güce ilişkin bilgiler bulunduğunu belirlemişlerdir. Farklı disiplinlerdeki çalışmaların istatistiksel güçleri düşük, orta ve yüksek etki büyüklüğünde farklılık göstermektedir. Gaskin ve Happell (2013) bir hemşirelik dergisinin 2010 ve 2011 yılı sayılarını incelemiş, düşük etki büyüklüğünde çalışmaların ortalama gücünün 0.34, orta etki büyüklüğünde 0.79, yüksek etki büyüklüğünde ise 0.94 olduğunu belirlemişlerdir. Okumura ve Sakamoto (2011) hemşirelik çalışmalarının da yer aldığı bir dergideki istatistiksel güçleri incelemiş, düşük, orta ve yüksek etki büyüklüklerinde çalışmaların ortalama güçlerinin sırasıyla 0.27, 0.71 ve 0.89 olduğunu belirlemişlerdir. Zhan (2013) İktisadi bir dergide 2003-2008 yılları arasındaki sayıları incelemiş ortalama istatistiksel güçlerinin sırasıyla 0.49, 0.93 ve 0.99 olduğunu belirlemiştir.

Literatür incelendiğinde standart etki büyüklüğü aralıklarını kullanarak farklı disiplinlere ait istatistiksel güçleri hesaplamının mümkün olduğu görülmektedir. Bu hesaplama çalışmalara alınan örneklem sayılarının yeterliliği ve sonuçlarının güvenilirliği hakkında bilgi verecektir. Bu araştırmanın amacı bir hemşirelik dergisinde 2009-2011 yılları arasında yayınlanmış olan makalelerde kullanılan istatistiksel testlerin güç analizini yapmak ve sunmaktır.

GEREÇ VE YÖNTEM

Tanımlayıcı türde yapılan bu çalışmada bir hemşirelik dergisinde 2009-2011 yılları arasında yayınlanmış araştırma makaleleri incelenerek istatistiksel güçleri hesaplanmıştır. Dergiye ait makalelere derginin web sitesinden ulaşılarak, gerekli veriler Ekim-Aralık 2011 tarihleri arasında toplanmıştır.

Çalışmada incelenen hemşirelik dergisi, yılda 4 sayı olarak çıkarılmaktadır. On-line olarak 1998 yılı sayılarından itibaren erişilebilen dergide 2009-2011 yılları arasında toplam 118 makale yayınlanmıştır (Tablo 2). Bu yayınlardan araştırma kriterlerine uygun 61 araştırma makalesi incelenmiş ve makalelerde kullanılan 725 istatistiksel testin güç analizi yapılmıştır.

İstatistiksel çözümleme içermeyen araştırma makaleleri, derlemeler, ölçek geliştirme veya adaptasyon makaleleri incelemeye alınmamıştır. Standart dışı tablo kullanımı, test sonucunun tabloda verilmemesi ve standart sapma, ortalama, örneklem sayısı verilmeyen 16 istatistiksel test kapsam dışında bırakılmıştır. Ayrıca 5 korelasyon tablosu ve 4 regresyon modeli değerlendirilmemiştir.

Tablo 2. Derginin Yıllara Göre İçeriği

Makale Türü	Yıllar			Toplam
	2009	2010	2011*	
Araştırma Makalesi (İstatistiksel Çözümleme Var)	30	18	13**	61
Araştırma Makalesi (İstatistiksel Çözümleme Yok)	3	9	11	23
Derleme	11	10	5	26
Ölçek Geliştirme veya Adaptasyonu	5	3	0	8
Toplam Makale Sayısı	49	40	29	118

* Derginin 2011 yılı son sayısı çıkmadığı için 3 sayı olarak değerlendirilmiştir.

**2011 yılında bir makaleye on-line ulaşılmadığı için bu yayın istatistiksel çözümleme yok kategorisine kabul edilmiştir.

Veri Toplama Araçları

Veriler, araştırmacı tarafından oluşturulan "Yayın tanıtım formu" aracılığıyla toplanmıştır. Formda toplam 10 soru bulunmaktadır (Yayın türü, istatistiksel güç hesaplanma durumu, yayın yılı ve sayısı, istatistiksel testlerin kullanılma adedi, kullanılan testin türü, testlerde hesaplanan örneklem sayısı, yüzde, ortalama, standart sapma vs.). On-line olarak taranan makaleler tek tek incelenerek yayın tanıtım formuna kaydedilmiştir.

Süreç

İstatistiksel gücün hesaplanabilmesi için örneklem sayısı, önemlilik düzeyi ve etki büyüklüğünün bilinmesine gerekir (Gravetter and Wallnau 2009). İncelenen makalelerde örneklem sayısı makalenin yöntem kısmından veya tablolardan bakılarak elde edilmiştir. Önemlilik düzeyi ise makale içinde aksi iddia edilmedikçe anlamlılık düzeyi, $\alpha=0.05$ olarak alınmış ve iki yönlü test olarak kabul edilmiştir. Etki büyüklüğünü belirlemek için Cohen (1988) tarafından tanımlanan standart düşük, orta ve yüksek etki büyüklükleri kullanılmıştır (Tablo 1). Örneğin örneklem sayısı ve önemlilik düzeyi bilinen bir t testi karşılaştırmasında standart olan düşük, orta ve yüksek etki büyüğü değerleri yerine konularak ilgili t testinin gücü her üç düzeyde de hesaplanmıştır (Cohen 1988). Bu nedenle bir istatistiksel karşılaştırmanın test gücü hesaplanırken üç farklı sonuç elde edilmiştir.

İstatistiksel bir testin gücü farklı etki büyüklüğü düzeylerinde farklılık göstermektedir. T testinin kullanıldığı bir istatistiksel karşılaştırmada düşük etki büyüklüğü ile hesaplanınca test gücü 0.30 çıkabilmektedir. Fakat aynı karşılaştırmanın gücü orta ve yüksek etki büyüklüklerinde 0.50 veya 0.80 çıkabilir (Zhan 2013). Cohen tarafından bildirilen standart etki büyüklüklerinin kullanılmasındaki amaç, bu çalışmanın sonuçlarının diğer çalışmalar ile kıyaslanabilmesini sağlamaktır (Brock 2003; Gaskin and Happell 2013; Okumura and Sakamoto 2011; Zhan 2013). İncelemeye alınan yayınlarda istatistiksel testlerin dağılımı ve bu

testlerin düşük, orta ve yüksek etki büyüklüklerindeki güçleri tablo şekline dönüştürülmüş ve diğer çalışmalar ile karşılaştırılmıştır. Çalışmada, Cohen (1988) tarafından standart etki büyüklüğü değerleri bildirilen bağımsız gruplarda t testi, χ^2 ve varyans analizinin yanı sıra, Mann-Whitney U, Wilcoxon, Kruskal-Wallis testlerinin de güçleri incelenmiştir. Non-parametrik testlerin güçleri, eş değeri olan parametrik testin standart etki büyüklükleri kullanılarak hesaplanmıştır (Parajapati, Dunne and Armstrong 2010; Smith, Hardy, Gammell 2011).

Çalışmaya başlamadan önce dergi editörlüğünden izin alınmıştır. Verilerin sunumunda sayı ve yüzdeler, mod ve medyan kullanılmıştır. Dergideki makalelerde kullanılan testlerin gücü G*Power paket programı ile hesaplanmıştır, sayısal veriler ise SPSS 11 paket programına aktarılarak incelenmiştir.

BULGULAR VE TARTIŞMA

Çalışmada 210 t testi, 160 varyans analizi, 76 χ^2 testi, 93 Mann-Whitney U testi, 54 Wilcoxon testi ve 132 Kruskal-Wallis testi incelenmiştir. T testi için düşük etki büyüklüğünde ortalama istatistiksel güç 0.34 ± 0.21 , orta etki büyüklüğünde 0.80 ± 0.18 , yüksek etki büyüklüğünde 0.97 ± 0.06 olarak hesaplanmıştır. Varyans analizi için ise, düşük etki büyüklüğünde ortalama istatistiksel güç 0.30 ± 0.13 , orta etki büyüklüğünde 0.88 ± 0.18 , yüksek etki büyüklüğünde 0.97 ± 0.50 olarak hesaplanmıştır. Diğer non-parametrik testler için ise istatistiksel güçler tablo 3'de görülmektedir.

Tablo 3. Kullanılan Testler ve İstatistiksel Güçlerinin Dağılımı

Kullanılan test	n		Düşük etki büyüklüğünde güç	Orta etki büyüklüğünde güç	Yüksek etki büyüklüğünde güç
t testi	210	Ort (ss)	0.34 ± 0.21	0.80 ± 0.18	0.97 ± 0.06
		Ortanca	0.30	0.85	0.99
Varyans analizi	160	Ort (ss)	0.30 ± 0.13	0.88 ± 0.18	0.97 ± 0.50
		Ortanca	0.31	0.98	0.99
χ^2	76	Ort (ss)	0.18 ± 0.95	0.73 ± 0.28	0.91 ± 0.15
		Ortanca	0.20	0.88	0.99
Mann-Whitney U	93	Ort (ss)	0.26 ± 0.21	0.64 ± 0.27	0.86 ± 0.20
		Ortanca	0.18	0.66	0.97
Wilcoxon	54	Ort (ss)	0.14 ± 0.12	0.39 ± 0.25	0.64 ± 0.29
		Ortanca	0.11	0.31	0.64
Kruskall-Wallis	132	Ort (ss)	0.15 ± 0.05	0.66 ± 0.15	0.95 ± 0.05
		Ortanca	0.14	0.68	0.97

Çalışmada testlere ait istatistiksel güçler için diğer yayınlar ile kıyaslanabilmesini sağlamak

değerleri üzerinden değerlendirilmiştir. Fakat analiz değerleri genel

olarak normal dağılım göstermemektedir ve ortalamalar yanıltıcı olabilir. Ortanca bir veri setinde tam ortada yer alan değerdir. Ortanca uç değerlere karşı hassas olmadığından özellikle normal dağılmayan verilerde kullanılabilir (Kalaycı 2010). Bu nedenle normal dağılmayan verilerde daha güvenli sonuç verebilen ortanca değerlerde çalışmada sunulmuştur. Çalışmalarda genellikle önemlilik düzeyi, p değeri ve

kullanılan istatistiksel analizler verilirken testin gücü hakkında bilgi verilmemektedir (Olinsky, Schumacher and Quinn 2012). Bu çalışmada da istatistiksel çözümleme kullanılan 61 çalışmanın yalnızca bir tanesinde istatistiksel güç hesaplanmış ve değeri makale içinde bildirilmiştir. Bu durum hemşirelik çalışmalarında istatistiksel güce yeterli önemin verildiğinin bir göstergesi olabilir.

Tablo 4. Test Türlerine Göre İstatistiksel Güçlerin Dağılımı

Testin Türü	n		Düşük Etki Büyükliğünde Güç	Orta Etki Büyükliğünde Güç	Yüksek Etki Büyükliğünde Güç
Parametrik (t, F)	370	Ort (ss)	0.32±0.18	0.83±0.19	0.97±0.53
		Ortanca	0.31	0.92	0.99
Non – Parametrik (X², U, Z, KW)	355	Ort (ss)	0.18±0.14	0.63±0.25	0.87±0.20
		Ortanca	0.14	0.66	0.97
Tüm testler (t, F, X², U, Z, KW)	725	Ort (ss)	0.25±0.18	0.73±0.24	0.92±0.15
		Ortanca	0.19	0.76	0.99

İstatistiksel testler parametrik ve non-parametrik olarak gruplandırılıp incelendiğinde; parametrik testler için (n=370), düşük etki büyüklüğünde istatistiksel gücün ortalamasının 0.32±0.18, orta etki büyüklüğünde 0.83±0.19, yüksek etki büyüklüğünde 0.97±0.53 olduğu görülmektedir. Non-parametrik testler için ise (n=335), düşük etki büyüklüğünde istatistiksel gücün ortalamasının 0.18±0.14, orta etki büyüklüğünde 0.63±0.25, yüksek etki büyüklüğünde 0.87±0.20 olduğu görülmektedir. Tablo 3 ve 4'teki bulgular incelendiğinde, tüm etki büyüklüğü sınıflamalarında parametrik testlerin daha yüksek güçte olduğu görülmektedir. Zodyey (2004), bir testin örneklem hacminin gücü ile doğru orantılı olduğunu bildirmektedir. Bu çalışmada parametrik testlerin daha güçlü çıkmasının sebebi daha büyük örneklerde kullanılmış olmalarından kaynaklanabilir.

Tüm testler birlikte incelendiğinde, makalelerde kullanılan testlerin düşük etki büyüklüğünde ortalama gücünün 0.25±0.18, orta etki büyüklüğünde 0.73±0.24, yüksek etki büyüklüğünde 0.92±0.15 olduğu görülmektedir (Tablo 4). Literatür taramasında son 10 yılda çok az sayıda benzer makaleye ulaşılmıştır. Daha eski makalelere ise güncelliğini kaybettiğinden dolayı bu çalışmada değinilmemiştir. Bu çalışma ile benzer çalışmalara bakıldığında, Gaskin ve

Happell (2013) bir hemşirelik dergisinin 2010 ve 2011 yılı sayılarını incelemiş, düşük etki büyüklüğünde çalışmaların ortalama gücünün 0.34, orta etki büyüklüğünde 0.79, Yüksek etki büyüklüğünde ise 0.94 olduğunu belirlemişlerdir. Okumura ve Sakamoto (2011) hemşirelik çalışmalarının da yer aldığı bir dergideki istatistiksel güçleri incelemiş, düşük, orta ve yüksek etki büyüklüklerinde çalışmaların ortalama güçlerinin sırasıyla 0.27, 0.71 ve 0.89 olduğunu belirlemişlerdir. Diğer disiplinlerde yapılan çalışmalarda ise, Zhan (2013) iktisadi bir dergide 2003-2008 yılları arasındaki sayıları incelemiş ortalama istatistiksel güçlerinin sırasıyla 0.49, 0.93 ve 0.99 olduğunu belirlemiştir. Brock (2003) pazarlama alanında düşük etki büyüklüğünde makalelerin ortalama istatistiksel güçlerini 0.29, orta etki büyüklüğünde 0.77 ve yüksek etki büyüklüğünde 0.93 olarak saptamıştır. Birçok disipline ait çalışmaların yayınlandığı dergilerde istatistiksel güçlerin değişkenlik gösterdiği görülmektedir. Çalışmalarda istatistiksel gücün 0.80 ve üzerinde olması istenir (Cohen 1988; Murphy and Myors 2004). Bizim çalışmamızdaki bulgulara göre özellikle düşük etki büyüklüğünde testlerin sadece %1.7'si bu gücü karşılamakta, orta ve yüksek etki büyüklüğünde ise, sırası ile %46 ve %88.7'si bu gücü karşılamaktadır (Tablo 5).

Tablo 5. Düşük, Orta ve Yüksek Etki Büyüklüğüne Göre İstatistiksel Güçlerin Dağılımı

GÜÇ	Düşük Etki Büyüklüğü		Orta Etki Büyüklüğü		Yüksek Etki Büyüklüğü	
	n	%	n	%	n	%
≤ 0.25	468	64.6	33	4.6	10	1.4
0.26 – 0.49	184	25.4	99	13.7	16	2.2
0.50 – 0.79	61	8.4	257	35.4	56	7.7
0.80 – 0.96	12	1.7	136	18.8	163	22.5
≥ 0.97	-	-	200	27.6	480	66.2

Genel anlamda düşük etki büyüklüğü, ortalamalar arasındaki farkın az olduğu anlamına gelir (Cohen 1988; Lan ve Lian 2010; Zodpey 2004). Benzer şekilde orta ve yüksek etki büyüklüklerinin de ortalamalar arasındaki farkın orta düzeyde veya büyük düzeyde olduğu anlamındadır. Bu anlamda incelenen çalışmalarda etki büyüklüklerinin orta düzeyde olduğu varsayılırsa %46'sının, büyük olduğu varsayılırsa %88.7'sinin istenilen güçte olabileceği görülmektedir. Tüm bu bulgular bazı araştırmaların istenilen güce (en az 0.80) ulaşması için daha fazla örnekleme ihtiyaç duyduğunu göstermektedir.

SONUÇ VE ÖNERİLER

Çalışmanın sonuçlarına göre hemşirelik araştırmalarında güç analizi yaparak örneklem sayısını belirlemek yaygın değildir. Sadece bir

çalışmada güç analizi yapılarak metin içinde verilmiştir. Parametrik testlerin kullanıldığı durumlarda istatistiksel güçler daha yüksektir. İncelenen çalışmaların bir kısmının istenilen güce ulaşması için daha fazla örnekleme ihtiyacı vardır. Bu nedenle hemşire akademisyenlere araştırmalarına başlamadan önce istatistiksel güç hesaplayarak örneklem büyüklüğünü belirlemeleri önerilmektedir. Özellikle deneysel çalışmalarda, yapılacak girişimin az düzeyde etki sağlayacağı ön görülüyorsa 0.80 ve üzeri istatistiksel gücü yakalamak için örneklem sayısının fazla tutulması önerilmektedir. Çalışmada yalnızca bir hemşirelik dergisinin üç yıllık yayınları incelenmiştir. Sonuçlar değerlendirilirken bu sınırlılık dikkate alınmalıdır.

KAYNAKLAR

- Brock JKU.** The Power of International Business Research. JIBS 2003; 34: 90-9.
- Cohen J.** Statistical power analysis for the Behavioral Sciences. New Jersey: Lawrence Erlbaum Associates, Inc.; 1988. p.20-7.
- Coolican H.** Research Methods and Statistics in Psychology. New York: Routledge; 2013, p.338-41.
- Demirel N, Gürler S.** Klinik çalışmalarda örneklem genişliğinin belirlenmesine pratik yaklaşımlar. Kafkas Univ Vet Fak Derg 2010; 16 (2): 205-11.
- Dorey F.** Statistics in brief: Statistical power: what is it and when should it be used?. CORR 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. Int J Ment Health Nurs 2013; 22 (1): 69-75.
- Gravetter VJ, Wallnau LB.** Statistics for the behavioral sciences. Belmont: Wadsworth Cengage Learning, 2009, p.265.
- Healey JF.** Statistics: A Tool for Social Research. Belmont: Wadsworth Cengage Learning; 2009, p.156-63.
- Kalaycı Ş.** SPSS uygulamalı çok değişkenli istatistik teknikleri. Ankara: Asil Yayınevi; 2010, p.320-50.
- Lan L, Lian Z.** Application of Statistical Power Analysis - How to Determine the Right Sample Size in Human Health, Comfort and Productivity Research. Building and Environment 2010; 45: 1202-13.

- Murphy KR, Myors B.** Statistical power analysis, a simple and general model for traditional and modern hypothesis test. London, Lawrence Erlbaum Associates, 2004, p.418-22.
- Olinsky A, Schumacher P, Quinn J.** The importance of teaching power in statistical hypothesis testing. IJMTL 2012; Dec:387-404.
- Okumura Y, Sakamoto S.** Statistical power and effect sizes of depression research in Japan. Psychiatry Clin Neurosci 2011; 65 (4): 356-64.
- Prajapati B, Dunne M, Armstrong R.** Sample size estimation and statistical power analyses. Optometry Today 2010; (July): 1-9.
- Smith DR, Hardy ICV, Gammell MP.** Power rangers: no improvement in the statistical power of analyses published in Animal Behaviour. Animal Behaviour 2011; 81 (1): 347-52.
- Usami S.** Statistical power of experimental research with hierarchical data. Behaviormetrika 2011; 38 (1): 63-84.
- Wang LL.** Retrospective statistical power: fallacies and recommendations. Newborn and Infant Nursing Reviews 2010; 10 (1): 55-9.
- Zhan G.** Statistical power in international business research: Study levels and data types. IBR 2013; 22 (4): 678-86.
- Zodpey SP.** Sample Size and Power Analysis in Medical Research. Indian J Dermatol Venereol Lepro 2004; 170 (2): 123-8.