

Etki Büyüklüğü ve Klinik Önemlilik

Sıddık Keskin^{1*}

¹*Biyoistatistik Anabilim Dalı, Temel Tıp Bilimleri, Bölümü / Van Yüzyüncü Yıl Üniversitesi Tıp Fakültesi, Türkiye*

skeskin@yyu.edu.tr

Özet – Bilimsel araştırmalarda, gruplar (uygulama, muamele, müdahale) arası farkın veya özellikler (değişkenler) arası ilişkinin, tesadüften ileri gelip gelmediğini belirlemek üzere, ifade edilen test hipotezine dayalı istatistik önemlilik (anlamlılık) testi yapılır. Hipotez testi ile; farkın veya ilişkinin tesadüften ileri gelmediği sonucuna varılmış ise diğer bir ifadeyle, istatistik olarak önemli ilişki veya fark bulunmuşsa, bu etkinin (veya ilişkinin) ne kadar büyük olduğu ve klinik (veya pratik) olarak da önemli (anlamlı) olup olmadığı veya uygulanabilmesi için yeterince büyük olup olmadığı sorgulanır. Klinik önemlilik, herhangi bir ilacın, tedavinin veya muamelenin pratikteki değerine veya risk, maliyet, hasta tercihi gibi kavramlar dikkate alınarak pratikte uygulanabilirliğine dayalı bir kavramdır. Bu durum, kısaca klinik (veya eşdeğer olarak pratik) önemlilik (anlamlılık) olarak ifade edilir. Klinik önemliliğe karar vermede, istatistik önemlilik testi gibi doğrudan bir (klinik) önemlilik testi yoktur. Ancak, klinik önemliliğe karar vermeye yardımcı olan ölçütlerden birisi de etki büyüklüğü (effect size) dür. Etki büyüklüğü, yapılan bir araştırmada, deneme sonucundaki gözlenen etkinin büyüklüğünün rakamsal (kantitatif) ölçüsüdür ve uygulanan tedavinin (muamelenin veya müdahalenin) ‘küçük’, ‘orta’ veya ‘büyük’ etkili olup olmadığı hakkında bilgi verir. Klinik önemliliğe ve yorumlanmasına katkı sağlamak üzere, r ailesi etki büyüklüğü ölçüleri, (r family effect size measures) d ailesi etki büyüklüğü ölçüleri (d family effect size measures) ve risk etkisi (risk potency) olmak üzere üç tip etki büyüklüğü önerilmektedir. Bu çalışmada etki büyüklüğü kavramı açıklanarak, yalnızca d ailesi etki büyüklüğü ölçülerine değinilmiş ve klinik önemlilik ile olan ilişkisi incelenmiştir.

Anahtar Kelimeler – Etki Büyüklüğü, Klinik Önemlilik, Örneklem Büyüklüğü, Meta Analizi, Cohen’s D

I. GİRİŞ

Bilimsel çalışmalarda, özellikle, klinik araştırmalarda ve Davranış Bilimleri ile ilgili çalışmalarda, genel olarak ilgilenilen değişkenler arasında ilişki olup olmadığı veya uygulamanın (tedavinin, müdahalenin veya muamelenin) kontrol grubundan farklı olup olmadığı ya da (farklı) uygulamalar arasında fark olup olmadığı incelenir. Bunun için ilk olarak gözlenen bu farkın veya ilişkinin, tesadüften ileri gelip gelmediğini belirlemek üzere, ifade edilen hipoteze dayalı istatistik önemlilik (anlamlılık) testi yapılır.

Hipotez testi ile farkın veya ilişkinin tesadüften ileri gelmediği sonucuna varılmış ise diğer bir ifadeyle, istatistik olarak önemli ilişki veya fark elde edilmişse, bu etkinin (veya ilişkinin) ne kadar büyük

olduğu, klinik (veya pratik) olarak da önemli (anlamlı) olup olmadığı veya uygulanabilmesi için yeterince büyük olup olmadığı sorgulanır. Bu durum, kısaca klinik (veya eşdeğer olarak pratik) önemlilik (anlamlılık) olarak ifade edilir.

Klinik önemliliğe karar vermede, istatistik önemlilik testi gibi doğrudan bir klinik önemlilik testi yoktur. Klinik önemliliğe karar vermeye yardımcı olan ölçütlerden birisi de etki büyüklüğü (effect size) dür. Diğer bir ifade ile klinik çalışmalarda, klinik ya da pratik önemliliği belirlemek üzere etki büyüklüğünden yararlanılmaktadır.

Bilimsel çalışmalarda etki büyüklüğünün önemi gittikçe artmaktadır. Zira birçok dergi, etki büyüklüğünün çalışmaya dahil edilmesini

istemektedir [1]. Bu nedenle çalışmada, etki büyüklüğü kavramı açıklanarak, klinik önemlilik ile olan ilişkisine değinilmiştir.

II. MATERYAL VE YÖNTEM

Çalışmada, etki büyüklüğü kavramı ve klinik önemlilikle olan ilişkisi açıklanmıştır.

A. Etki Büyüklüğü (Effect Size)

‘Etki büyüklüğü’ kavramını oluşturan etki, herhangi bir faktörün veya değişkenin, diğer bir değişken (özellik) üzerine kantitatif (nümerik, sayısal veya rakamsal) yansımaları olarak ifade edilirken, büyüklük ise bu yansımaların (kantitatif rakamsal veya nümerik) miktarıdır. Böylece etki büyüklüğü, ilgilenilen herhangi bir özelliğe, etkili olabileceği düşünülen veya bu özellik ile ilişkili olabileceği varsayılan diğer bir özelliğin büyüklüğünün nicel (kantitatif) yansımaları olarak tanımlanır [2].

Etki büyüklüğü, biyomedikal araştırmalarda ve özellikle randomize klinik çalışmalarda; örneklem büyüğü ve güç değerinin hesaplanmasında kullanıldığı gibi Meta analizlerinde de genel etki büyüklüğünün hesaplanmasında kullanılmaktadır.

Araştırma sonuçlarının önemini artırmak için pratik öneme sahip göstergelerden birisi ve belki de en önemlisi etki büyüklüğüdür. Etki büyüklüğü, bir uygulama veya muamele etkisi için, yapılacak istatistik hipotez testi sonucunda, istatistik olarak önemli (anlamlı) bulunması öngörülen değerdir [3]. Etki büyüklüğü, bir tedavinin etkisinin yönünü ve büyüklüğünü, iki tedavi grubu arasındaki farkı veya herhangi bir sayısal karşılaştırmayı kolay yorumlanabilir sağlayan bir değerdir.

Klinik önemliliğe ve yorumlanmasına katkı sağlamak üzere, r ailesi etki büyüklüğü ölçüleri, (r family effect size measures) d ailesi etki büyüklüğü ölçüleri (d family effect size measures) ve risk etkisi (risk potency) olmak üzere üç tip etki büyüklüğü önerilmektedir [4]. Bu çalışmada yalnızca d ailesi etki büyüklüğü ölçülerine değinilmiştir.

İki grup ortalaması için en yaygın kullanılan etki büyüklüğü ‘Cohen’s d’ olarak ifade edilen etki büyüklüğüdür. Bu değer, Bağımsız gruplar t testinde ve varyans analizinde (ANOVA) kullanılabilir.

Cohen’s d değeri

Cohen’s d değeri olarak Etki büyüklüğü (d);

$$d = \frac{\bar{X}_1 - \bar{X}_2}{s} \quad (1)$$

eşitliği ile hesaplanır. (1) no’lu eşitlikte; ‘S’, Toplanmış standart sapma (pooled standard deviation) olup;

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

olarak hesaplanır [5].

(2) no’lu eşitlikte; n_1 ve n_2 sırasıyla, birinci ve ikinci gruba ait örnek genişliği iken; S_1^2 ve S_2^2 de sırasıyla birinci ve ikinci gruba ait varyanstır. (1) no’lu eşitlikteki Cohen’s d, iki grup arasındaki farklılığı, standart sapma cinsinden belirtir. Örneğin; iki grup arasındaki d değeri 1 bulunmuş ise iki grup arasında 1 standart sapmalılık fark (uzaklık veya açıklık) vardır demektir.

(1) no’lu eşitlikte bulunan d değeri; 0.2 ise ‘küçük etki büyüklüğü’, 0.5 ise ‘orta etki büyüklüğü’, 0.8 ve daha büyük ise ‘yüksek etki büyüklüğü’ olarak ifade edilir. Etki büyüklüğünün, 0.2 ve daha küçük olması, sonuç istatistik olarak önemli olsa bile, farkın klinik olarak önemli olmadığı şeklinde ifade edilir.

Hedge’s g değeri

$$\text{Hedge’s g değeri; } g = \frac{\bar{X}_1 - \bar{X}_2}{S^*} \quad (3)$$

olarak hesaplanır.

$$(3) \text{ no’lu eşitlikte ‘S*’; } S^* = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \quad (4)$$

olarak hesaplanır [5].

Cohen’s d değerini hesaplamada toplanmış standart sapma kullanılırken, Hedge’s g değerinin hesaplanmasında ağırlıklandırılmış toplanmış standart sapma kullanılır [5].

Her iki etki büyüklüğü de popülasyon etki büyüklüğü için pozitif yanlılığa sahiptir. Zira standart sapmanın örnekten hesaplanmış olması, gerçek popülasyon standart sapmasından daha küçük standart sapma elde edilmesi eğilimi gösterdiği için pozitif yanlılık söz konusu olabilmektedir. Ancak, büyük örneklerde bu yanlılık göz ardı edilebilir [5].

Diğer yandan bu yanlılığı düzeltmek üzere Hedge’s g* değerinin hesaplanması önermiştir.

Hedge’s g*; $g^* = J(n_1 + n_2 - 2)g$ olarak hesaplanmakta olup eşitlikteki J değeri;

$$J(a) = \frac{\Gamma(\frac{a}{2})}{\sqrt{\frac{a}{2}}\Gamma(a/2)} \quad (5)$$

olarak hesaplanır. (5) no’lu eşitlikte Γ ; Gama fonksiyonudur [5].

Glass's Δ değeri

$$\text{Glass's } \Delta \text{ değeri } \Delta = \frac{\bar{X}_1 - \bar{X}_2}{S_2} \quad (6)$$

olarak hesaplanır. (6) no'lu eşitlikteki S_2 değeri, Standart veya Kontrol grubuna ait standart sapmadır [5].

B. Klinik Önemlilik (Anlamlılık)

Klinik önemlilik, herhangi bir ilacın, tedavinin veya muamelenin pratikteki değerine veya risk, maliyet, hasta tercihi gibi kavramlar dikkate alınarak pratikte uygulanabilirliğine dayalı bir kavramdır. Diğer bir ifade ile klinik önemlilik, uygulanan herhangi bir tedavinin etkisinin, pratikte önemi, uygulanabilirliği ve hasta üzerinde gerçek (elle tutulur, gözle görülür veya belirgin) bir etkisinin olmasıdır. Diğer yandan klinik olarak önemli bulgular, hastayı tıbbi bakımdan iyileştiren, bireyin fiziksel işlevini, zihinsel durumunu ve sosyal yaşama katılma yeteneğini belirgin geliştiren bulgulardır.

III. TARTIŞMA

Başta klinik araştırmalar olmak üzere, birçok çalışmada, etki büyüklüğü önemli bir kavramdır. Zira araştırmacılar, herhangi bir müdahalenin, tedavinin, uygulamanın, muamelenin veya deneysel bir işlemin etkisinin, sıfırdan farklı olup olmadığını veya büyüklüğünü bilmek (öğrenmek) ister. Bu durumda etki büyüklüğü kullanılabilir. Kelley ve Preacher [2], etki büyüklüğünün 10 farklı yönünden bahsederek, bunlardan birisinin de; pratik, klinik, tıbbi veya yönetsel (practical, clinical, medical, or managerial) önemlilik hakkında bilgi verdiğini, ancak etki büyüklüğünün belirli bir değerine atfedilen önem düzeyinin, bir alandan diğerine büyük ölçüde değişebileceğın vurgulamıştır.

Araştırmacı veya uzman hekim görüşü ile klinik önemliliğe karar vermede; uygulamanın (veya tedavinin) yararı, maliyeti, riski, ya da etkileri gibi faktörler dikkate alınır. Zira, klinik önemliliğe karar vermede, istatistik önemliliğin testi gibi bir klinik önemlilik testi yoktur. Ancak, klinik önemliliğe ve yorumlanmasına katkı sağlamak üzere, r ailesi etki büyüklüğü ölçüleri, (r family effect size measures) d ailesi etki büyüklüğü ölçüleri (d family effect size measures) ve risk etkisi (risk potency) olmak üzere üç tip etki büyüklüğü önerilmektedir [4].

Etki büyüklüğü, yapılan bir araştırmada, deneme sonucundaki gözlenen etkinin büyüklüğünün

rakamsal (kantitatif) ölçüsüdür. Araştırmacı, herhangi bir uygulamanın, müdahalenin veya muamelenin yapıldığı bir deneysel grup ile kontrol grubunu karşılaştırmak isteyebilir. Etki büyüklüğü, bu durumda, uygulanan tedavinin (muamelenin veya müdahalenin) küçük, orta veya büyük etkili olup olmadığı hakkında bilgi verir.

Etki büyüklüğü, standardize edilmiş bir değer olarak, cevap (çıktı) değişkeninin veya bağımlı değişkenin ölçü biriminden bağımsız bir metrik ölçüdür. Bu nedenle etki büyüklüğü, sonuçların klinik veya pratik önemliliği hakkında araştırmacılara bilgi sağlar. Oysaki istatistik önemlilik, yalnız başına denemeden gözlenen sonuçların, ne kadar olduğu popülasyonda ne kadar etkili olabileceği hakkında bilgi vermeyebilir.

Meta analizlerinde, standardize edilmiş etki büyüklüğü ile farklı çalışmalardaki muamele (müdahale, tedavi veya uygulama) etkileri karşılaştırılabilir veya yorumlanabilir. Ayrıca, önceki çalışmalardan elde edilen etki büyüklüğü, yeni çalışmaların planlanmasında, güç analizinde ve çalışmadan istatistik olarak önemli sonuç elde edilebilmesi için gerekli olan asgari örneklem genişliğinin belirlenmesinde kullanılmaktadır [6].

Etki büyüklüğü ile ilgili belki de en önemli sorun, özellikle, örneklem genişliği hesaplamada, etki büyüklüğünün ne kadar alınacağı veya nasıl belirleneceğidir. Zira etki büyüklüğünün düşük olarak alınması, örneklem büyüklüğünü artırır ve popülasyonda gerçekte önemli bir farklılık yokken, araştırmacıların klinik veya pratik önemliliği destekleyecek istatistik önemliliği elde etmelerini sağlar. Bunun için Pogrow [1] klinik veya pratik önemliliği destekleyecek istatistik önemliliği elde etmek üzere; gerekli minimum örneklem genişliği için etki büyüklüğünün '0.2' olarak alınabileceğini belirtmiştir.

Citrome [8], sonuçların klinik önemini ölçmek için klinik araştırma sonuçlarına ait raporların, etki büyüklüğü ölçümlerini de içermesi gerektiğini vurgulayarak, p değerinin tedavi etkilerinin büyüklüğü hakkında bilgi verici olmadığını belirtmiştir. Ayrıca, klinik önemlilik için Cohen's d istatistiği ve farklı versiyonlarının sıklıkla kullanıldığına, ancak günden güne değişen klinik pratiklerde uygulanabilirlik açısından anlaşılmasının kolay olmadığına dikkat çekmiştir.

IV. SONUÇLAR

Etki büyüklüğü, yapılan bir araştırmada, deneme sonucunda gözlenen etkinin büyüklüğünün, rakamsal (kantitatif) ölçüsüdür ve klinik önemliliği değerlendirmek için kullanılabilir. Bu nedenle, klinik veya pratik önemliliğin değerlendirilmesi veya yorumlanmasına katkı sağlayabilmesi açısından, özellikle klinik araştırmalarda, sonuçların raporlanmasında etki büyüklüğünün de verilmesinin önemli olduğu düşünülmektedir.

KAYNAKLAR

- [1] L. S. Peterson, " Clinical" Significance:" Clinical" Significance and" Practical" Significance are NOT the Same Things. ERIC Clearinghouse. 2008.
- [2] K. Kelley and K. J. Preacher, On effect size. *Psychological Methods*, 17 (2), 137–152. 2012.
- [3] B. Thompson. 2006. *Foundations of behavioral statistics: An insight-based approach*. New York: The Guilford Press.
- [4] H.C. Kraemer, G. A. Morgan, N. L. Leech, J. A. Gliner, J. J Vaske and R. J. Harmon. Measures of clinical significance. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(12), 1524-1529. 2003.
- [5] S. Yang and G. Berdine. Effect size. *The Southwest Respiratory and Critical Care Chronicles*, 9(40), 65-68. 2021
- [6] D. Lakens. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, 4, 863. 2013.
- [7] S. Pogrow. How Effect Size (Practical Significance) Misleads Clinical Practice: The Case for Switching to Practical Benefit to Assess Applied Research Findings, *The American Statistician*, 73:sup1, 223-234, 2019.
- [8] L. Citrome. Quantifying clinical relevance. *Innov Clin Neurosci*. May;11(5-6):26-30. 2014.