



Klinik Arařtırmalarda Etik-İstatistik İliřkisi

Doç. Dr. Mutlu Hayran
Hacettepe Üniversitesi Onkoloji Enstitüsü,
Preventif Onkoloji AD

Bu makalede, etik sorun oluşturabilecek tasarım ile ilgili problemlerden istatistikle en yakın ilişkili olanlar sıralanmıştır. Aşağıda bahsedilen uygunsuz tip-1 ya da tip-2 hata seçimi, örneklem büyüklüğü uygunsuzluğu, sonuç noktasının belirsizliği, yanlış yöntem seçimi gibi sorunların tamamı hiç özenilmeden hazırlanmış, planlaması yanlış bir çalışmaya işaret eder. Bu tür hatalar, etik kurullarca da tespit edildiğinde onay öncesi arařtırıcıdan düzeltilmesi istenen konular arasında yer alır.

Bir klinik araştırma kapsamında, istatistik ile etik arasındaki ilişkiyi anlayabilmek için bir araştırma düzeni içerisinde çalışmaya başlarken önceden kabul ettiğimiz hatalar olan tip 1 ve tip 2 hatayı anlamak gerekmektedir. Bu hatalar, çalışma düzenindeki bireylerin tüm dünyadaki çalışmadakilere benzer bireyleri temsil ettiği varsayımı ile yapılan çıkarımlarda ortaya çıkacak kaçınılmaz hatalardır. Tip-1 hata ve tip-2 hata aynı zamanda, sırasıyla, alfa ve beta hatası olarak da adlandırılmaktadır. Yapılan herhangi bir çalışmada varılan sonucun gerçek hayattaki olayı doğru temsil edip edememe durumuna göre tanımlanmaktadır. Tip-1 hata gerçekte olmayan bir durumun (iki grup arasında fark, iki olay arasında ilişki gibi) çalışma düzeni içerisinde var olarak gösterilmesi, tip-2 hata ise gerçekte var olan bir durumun çalışma düzeni içerisinde gösterilememesi olarak tanımlanabilir (Tablo-1).

Tablo 1: Tip-1 ve Tip-2 hataların ortaya çıkma şekilleri

		Gerçek Hayat	
		Fark ya da İlişki Var	Fark ya da İlişki Yok
Çalışma İçinde Bulunan Sonuç	Fark ya da İlişki Bulundu	Çalışma başarılı	Çalışmada Tip-1 hata yapılmış.
	Fark ya da İlişki Bulunmadı	Çalışmada Tip-2 hata yapılmış.	Çalışma başarılı.

Gerçek hayatta çalışma koşullarına uyan bireylerin tamamını çalışmaya almadan tip-1 ve tip-2 hatalarının sıfırlanması imkansızdır. Açıktır ki, biz çalışmamıza sonuçları genelleyeceğimiz bireylerden ne kadar fazlasını dahil edersek yanılma olasılığımız o kadar az olacaktır. Çalışmalarda daha protokol yazılırken bu çalışma içerisinde ne oranda tip-1 ve tip-2 hata yapılmasının kabullenildiği önceden kararlaştırılır ve buna göre ne kadar hastanın çalışmaya alınması gerektiği belirlenir. Doğru hasta sayısının tespiti için örneklem büyüklüğü hesapları kullanılır. Bir çalışmadaki tip-1 hatanın en fazla %5, tip 2 hatanın da en fazla %20 boyutlarında olması beklenmektedir.

Çalışmanın primer sonuç noktasının belirtilmemesi: Çalışmanın amacında birden fazla hedef yer aldığı durumda bunların hangisinin ana hedef olduğu ve bu ana hedefe konu olan hipotezin doğru ya da yanlış olduğuna dair tam bir cevap verilebilmesi için bireylerde yapılan hangi ölçüm ya da değerlendirmenin primer parametre olarak kullanılacağı belirtilmelidir. Örneğin, primer amaç bir ilaç ile ağrının azaltılıp azaltılmadığını bulmak ise hastaya ağrı olup olmadığını sormak, hastaya ağrı şiddetini sormak, hekimin muayene sonucu görüşünü kaydetmek, kullanılan toplam ilaç miktarını kaydetmek, ilaçsız geçen süreyi kaydetmek, 0-100 arasında görsel analog skalası ile bir değerlendirme yapmak gibi birçok seçenek

olabilir. Bu seçeneklerden hangisinin kullanılacağı spesifik olarak belirtilmelidir. “Çalışmada A tedavisiyle B tedavisine göre ağrının istatistiksel olarak anlamlı düzeyde azaltıldığı bulunmuştur” gibi bir çıkarım yapıp yapılamayacağı bu ağrı parametresinin uygunluğu ile değerlendirilecektir.

Örneklem büyüklüğü hesabı: Bir çalışma yapıldığında iki olay arasında ilişki olup olmadığını incelemek, karşılaştırılan gruplar arasında fark olup olmadığını bulmak, ya da bazı bilgilere bakarak daha sonra olacak olaylar veya daha zor ölçülebilen bilgiler ile ilgili tahminlerde bulunmak araştırmacının başlıca merakları arasında yer alır. Birincil ve ikincil amaçlar bu yaklaşımlardan en azından birini kullanmaktadır. Pozitif sonuç bulma denilen durum, “ilişki var”, “fark var”, ya da “X,Y,Z değerlerine bakarak T değişkeni önceden tahmin edilebilir” gibi sonuçlara ulaşabilme durumudur. Bunun dışındaki durumlarda birçok çalışmada araştırmacılar, çalışma sonu raporunda “ilişki bulamadık”, “fark bulamadık” gibi negatif sonuç bildirmek zorunda kalırlar. Hatta negatif sonuç bulduklarında bir yaygın yapamayacaklarını düşünerek sonuçlardan hayal kırıklığı ile bahsederler. Oysa doğru yapılmış bir örneklem büyüklüğü hesabı ve doğru sayıda bireyin katılması sağlanan bir çalışma, negatif sonuç bulunduğu araştırmacıya “fark bulamadık” değil “fark olmadığını bulduk” deme olanağını sağlayacaktır. İkinci söylem artık “fark bulamadık”

söylemindeki gibi bir belirsizlik ve acizlik göstermeyen, bilim dünyasına katkı yapan kendinden emin bir söylemdir. Örneklem büyüklüğü iyi belirlenmeden belirsiz bir çıkarım yapma ihtimalimiz olan bir çalışmada insanları denek olarak kullanmak etik olmayacaktır.

Doğru yapılmış örneklem büyüklüğü hesabının önemli olduğu bir diğer durum, klinik olarak anlamlı boyutta farkların istatistiksel olarak da anlamlı bulunmasını sağlamaktır. Örneğin, iki tedavi yaklaşımının %60 ve %80'lik tedavi oranları sağladığı bir senaryo düşünelim. Bu %20'lik fark aslında bir klinisyenin önemseyebileceği boyutta bir farktır. Ancak, bu çalışma her gruba araştırmacıların sık kullandığı bir örneklem büyüklüğü olan 30 hasta alınarak yapılırsa istatistiksel olarak iki grup arasında fark yok çıkarımı yapmak zorunda kalacaktır ($p=0,16$, Ki-kare testi). Bu senaryo için tedavi edilemeyen hastaların öldüklerini düşünelim. Yüzde 20'ye karşı yüzde 40 ölümün “tedavi yöntemleri arasında anlamlı fark olmayan büyüklükte bir etki” olarak yorumlanması ve bu iki tedavinin birinin diğerine üstün olmadığına karar verilmesi, son derece yanlış ve etik olmayan bir çıkarım olacaktır.

Bu örnekte gerçekte kaç hasta alınması gerektiği, orantıların karşılaştırılması için kullanılan formüllerden biri olan arksinüs transformasyonu ile hesaplandığında her bir gruba 81 kişinin alınması gerektiği bulunacaktır.

$$\begin{aligned} N &= \frac{1}{2} \{ (z_{1-\alpha/2} - z_{\beta}) / (\sin^{-1} \sqrt{p_1} - \sin^{-1} \sqrt{p_2}) \}^2 \\ &= \frac{1}{2} \{ (1,96 - (-0,84)) / (0,88 - 1,10) \}^2 \\ &= \frac{1}{2} (2,80 / -0,22)^2 \\ &= \frac{1}{2} (-12,72)^2 \\ &= 161,80 / 2 = 81 \end{aligned}$$

Görülmektedir ki ciddi bir bilimsel araştırma için bazı araştırmacıların sık kullandığı “her gruba yaklaşık 30 hasta alalım” yaklaşımı etik sorunlar arasında yer alan bir “düzensiz araştırma” (*sloppy research*) örneği olarak karşımıza çıkabilmektedir.

Eşdeğerlik: Yukarıda tip-1 ve tip-2 hata için %5 ve %20 sınırlarının bilimsel olarak kabul edilen sınırlar olduğunu belirtmiştik. Hatırlanacağı gibi %20'lik yanılma payı “gerçekte fark olduğu durumda çalışma içerisinde fark bulamama” ihtimali idi. Asıl amacı bir tedavinin diğerinden daha iyi olabileceğini göstermek değil de, en az diğeri kadar etkili olup

olmadığını incelemek olan bir çalışmada, örneğin bir eşdeğerlik çalışmasında, bu tarz bir hatanın %20'den daha düşük olması hedeflenir. Aksi takdirde daha az etkin ürünler, kolaylıkla diğerleri kadar etkinmiş gibi bulunabilecektir. Bu nedenle eşdeğerlik çalışmalarının daha özel bir istatistiksel yaklaşım ile planlanmış olması gerekmektedir.

Kullanılan analiz yönteminin uygunsuzluğu:

Yukarıda bahsedilen tip-1 ve tip-2 hatalar, çalışmadaki diğer tüm basamaklara dikkat edildiği halde ortaya çıkma ihtimali olan hatalardır. İstatistiksel yöntemler doğru kullanılmadan yapılacak “fark var”, “ilişki olmadığı bulundu” gibi çıkarımlar yanlış bir bilginin bilim dünyasına eklenmesine yol açacaktır. Burada sadece p değerlerinin doğru hesaplanması değil, istatistiksel farkların klinik farklarla da birleştirilerek yani etki büyüklüğü de göz önüne alınarak değerlendirmesi gerekmektedir. İyi yapılmış bir istatistiksel analize dayandırılmayan çıkarımlar, daha sonra gereksiz yere faydasız tedavilere hasta randomize edilmesine, gereksiz yere uzayan çalışmalara, yanlış yere terk edilen tedavi yöntemlerine ve hatalı bir tıbbi bakıma kadar uzayabilen süreçler oluşturma riski ortaya çıkaracak hatalı bilgileri “kanıt” olarak bilimsel literatüre ekleyebilecektir.

Unutulmamalıdır ki, eğer çalışma bir araştırma sorusuna kesin bir şekilde EVET ya da HAYIR diye cevap veremiyorsa, ya da daha sonra yapılacak çalışmalara o sırada bu konuda çalışanlar için yararlı ve gerekli bir bilgiyi DOĞRU olarak üretemiyorsa, böyle bir çalışmada insanları denek olarak kullanmak ETİKDEĞİLDİR.

Kaynaklar:

1. Dobson AJ, Gebski VJ. Sample Sizes for Comparing Two Independent Proportions Using the Continuity-Corrected Arc Sine Transformation. *The Statistician*. 1986;35(1):51-53.
2. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human use: ICH Harmonised Tripartite Guideline. *Statistical Principles For Clinical Trials* (E9) 1998, p:1-35.
3. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human use: ICH Harmonised Tripartite Guideline. *Guideline for Good Clinical Practice* (E6(R1)) 1998, p:32-33.
4. Kansu E, Ruacan S. Research ethics and scientific misconduct in biomedical research. *Acta Neurochir Suppl*. 2002;85:11-5.
5. Hayran M. Hastane enfeksiyonları araştırmalarında istatistiksel analiz. *Hastane İnfeksiyonları Dergisi* 2006;10(4):227-235.