

BİLİMSEL MAKALELERDE İSTATİSTİĞİN KULLANIMI, SUNUMU VE SIK YAPILAN HATALAR

Ergun KARAAĞAOĞLU

Makalelerde İstatistiğin Sunumu

Temel istatistik kavramların yeterince bilinmemesi ve yanlış kullanımı bilimsel literatürün yaklaşık yarısında en az bir istatistiksel hatanın olduğunu ortaya koymakta⁽¹⁾ ve bu da bilimsel süreci tehlikeye sokmaktadır. İstatistiksel hatanın en aza indirilmesi hatta ortadan kaldırılabilmesi için değişik çalışmalar yürütülmüş, dergilerde istatistiğin nasıl rapor edilmesi (sunumu) konusunda birçok yönerge hazırlanmıştır. Bu yönergeler, makalelerin *Materyal Metod, Sonuçlar ve Tartışma* bölümlerinde istatistiğin sunumu hakkında önemli konulara dikkat çekmektedir^(1,2).

1. Araştırma planlanırken bir istatistikçiye danışılmalıdır.
2. Alfa, anlamlılık düzeyi araştırmanın amaçları ile uyumlu seçilmelidir. Alışlagelmiş 0.05 değeri her zaman uygun olmayabilir.
3. Kullanılan istatistiksel yöntem referans verilerek belirtilmelidir.
4. Çoklu karşılaştırmalar denetlenmelidir.
5. Değişkenliğin ölçüsü olarak standart sapma kullanılmalıdır.
6. Kesin olmayışlık (belirsizlik) güven aralıkları ile belirtilmelidir.
7. p'nin kesin değeri verilmelidir.
8. Rakamlardaki ondalık hane sayısı, kullanılan değişkenin ölçüm hassaslığı ile uyumlu olmalıdır.
9. Sonuçlar güven aralığı yada kesin bir p değeri verilerek değerlendirilmelidir.
10. İstatistiksel anlamlılık ile klinik önemlilik birbirine karıştırılmamalıdır.

Prof. Dr. Ergun KARAAĞAOĞLU - Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik A.D - ANKARA

Klinik Denemelerin Raporlanmasında İstatistiğin Sunumu :

Klinik denemelerde başlangıç verilerinin, her hasta için (baseline data) randomizasyon öncesinde kayıt edilmesi gerekir. Bu veri, hastanın o andaki durumu (semptomlar ve değişik ölçüm değerleri), tıbbi geçmişi (daha önceki hastalıklar, tanıdan itibaren geçen süre) ve demografik bilgilerinden (yaş, cins ve diğer kişisel özellikler) oluşur. Klinik deneme sonuçları raporlanırken, bu başlangıç verilerinin üç ana kullanım alanı vardır;

- i) Alt grup analizi, hastaların sonuç değerleri denemeler arasında farklılık (yada farksızlık) gösterdiğinde, bunun tüm hastalar yada hasta grupları için aynı olup olmadığını, bir ya da daha fazla başlangıç değerine bağlı olup olmadığını araştırmak için yürütülür.
- ii) Ortak değişkene göre düzeltilmiş analiz, hastanın sonuç değişkeni ile ilişkili başlangıç değişkenlerini (özellikle bu başlangıç değişkenleri deneme grupları arasında dengeli dağılmadıysa) göz önüne alarak, deneme etkilerinin ve farklılıklarının kestirimlerini en doğru biçimde elde etmek için kullanılır.
- iii) Başlangıç değerlerine göre karşılaştırmalar, farklı hasta gruplarını tanımlamak ve deneme başlangıcında gruplarının ne derece benzer olduğunu göstermek amacıyla yapılır⁽³⁾.

Burada, bu üç başlık için raporlama sırasında dikkat edilmesi gereken temel istatistiksel işlemler ve dikkat edilmemesi durumunda ortaya çıkabilecek hatalara değinilecektir.

Alt Grup Analizleri:

Klinik denemeye dahil edilen hastalar homojen bir örneklem oluşturamazlar. Tedaviye verdikleri yanıt, farklı tedavilerin onlar üzerinde oluşturduğu farklı etkiler, hangi tedavinin hangi hasta (grubu) için en iyi olabileceği seçimini etkiler.

Alt grup analizi yaparken çeşitli zorluklarla karşılaşılabilir yada hata yapılabilir:

- a. Çoğu klinik deneme, deneme grupları arasında yanıt değişkenindeki tümel farklılığı belirleyebilecek yeterli güce sahiptir. Eğer gerçekten alt grup etkileri

varsa, bunlar, deneme yeterli büyüklükte hasta grubu üzerinde yürütülmediğinden belirlenemeyebilir.

b.Çok fazla sayıda başlangıç değişkeni varsa ve bu değişkenlere ilişkin hangi alt grupların yeni tedaviye daha olumlu (yada olumsuz) yanıt verdiği konusunda belirgin bir ön tanımlama yoksa, fazla sayıda alt grup analizi yapılabilir ve bu bir çok alt grup analizi içerisinden ilginç olanların seçimi gibi bir yanılığa düşülebilir.

c.Deneme etkisinin farklı alt gruplarda önemli olup olmadığını göstermenin etkili yolu, etkileşimi test etmektir. Her farklı alt grupta deneme etkisini ayrı, ayrı test etmek ve bu testlere ilişkin ayrı, ayrı p değerleri vermek yanıltıcı olabilir. Tümel olarak deneme etkisi önemli olduğunda, bazı alt gruplar için bu etkinin şansa yada yetersiz örneklem genişliğine bağlı olarak önemli, bazılarında da önemsiz bulunacağı unutulmamalıdır.

d.Deneme etkisinin varlığının (yada olmadığının) belirli bir hasta grubuna genelleştirilemeyeceği ve alt gruplar hakkında bazı sonuçların abartılı olarak ortaya konmuş olması olasılığına karşı önlem almak gerekir. Örneğin, alt gruplar arasında denemelerin karşılaştırılması amacıyla yaşam eğrileri kullanıldığında, alt grup etkisinin olabileceğini gösteren bir çalışmada ⁽⁴⁾, doğru yöntem olan etkileşim test edilmiş olsaydı bu farklılık bulunamayacaktı.

Ortak Değişkene Göre Düzeltme:

Ortak değişkene göre düzeltme yapmanın amacı, deneme grupları arasında dengesiz dağılmış, prognoz üzerinde etkili olabilecek başlangıç değişkenlerini göz önüne alarak, denemeler arası yansız ve istatistiksel olarak etkili karşılaştırmalar yapabilmektir. Ortak değişken olarak düşünülebilecek başlangıç değişkenleri:

- Sonucu belirleyebileceği adımsal değişken seçimi yöntemi ile bulunmuş,
- Gruplar arasında dengeli dağılmamış,
- Randomizasyonda tabakalama amacıyla kullanılmış değişkenlerdir. Denemeler arasındaki farklılığa ilişkin en uygun p değerini

bulmak, yanıt değişkeni üzerinde deneme etkisinin yansız bir kestirimini elde etmek, istatistiksel gücü ve deneme etkisinin kestiriminin doğruluğunu artırmak için ortak değişkene göre düzeltme gerekli olabilir.

Başlangıç Değerlerinin Karşılaştırılabilirliği:

Çoğu klinik deneme raporunda “Tablo 1”, birçok başlangıç değişkeninin dağılımlarının deneme grupları arasında karşılaştırılmasına ayrılmıştır. Bunun amacı;

- Çalışmaya dahil edilen hasta örnekleminin başlangıç özelliklerini tanımlamak,
- Dengeli (benzer) deneme grupları elde edildiğinde, randomizasyonun iyi yapıldığını göstermek,
- Ortak değişkene göre düzeltme yapmaya gerek kalmadan, sonuçların güvenle, yansız olarak elde edildiğini göstermek,
- Deneme grupları arasındaki rastlantısal olarak ortaya çıkmış olabilecek dengesizliği belirleyebilmektir.

İlk amaç, denemeye katılanların özelliklerini belirtmesi ve bunun sonucu olarak da deneme sonuçlarının kimlere genelleştirilebileceğini göstermesi açısından çok önemlidir. Fazla sayıda başlangıç değerine ilişkin tanımlayıcı istatistikleri değil, kilit değişken olarak adlandırılacak önemli değişkenlere ilişkin olanları vermek, sayı ve yüzdelerin hem tüm grup içerisinde, hem de deneme grupları içerisindeki yüzdesini vermek yerine, sadece birini tercih etmek ve sürekli değişkenler için ortalamalarla birlikte standart hataları değil, hasta grupları arasındaki değişkenliği daha iyi açıklayan standart sapmaları vermek daha doğru olur.

Başlangıç değişkenlerine ilişkin karşılaştırmalar için önemlilik testlerinin gereksiz kullanımından kaçınmak gerekir. Şansa bağlı olarak önemli farklılık bulma olasılığı, karşılaştırma sayısı arttıkça artacaktır. Başlangıç değişkenindeki farklılığın önemi, bu değişkenin sonuç değişkeni ile olan ilişkisine bağlıdır. İstatistiksel olarak farklılığı önemli olmadığı halde, sonuç değişkeni ile kuvvetli

ilişkili olan bir değişkenin deneme grupları arasında dengesiz dağılmış olması bir problem yaratabilirken, sonuç değişkeni ile çok ilgisi olmayan bir başlangıç değişkeninin gruplar arasında istatistiksel olarak önemli ölçüde farklılaşması bir sorun olmayabilir. Bu nedenle, başlangıç değerlerinin karşılaştırılması için verilen p değerleri, doğru amaca hizmet etmezler. Bu tür tablolarda p değerlerine yer verilmemesi daha doğrudur.

Kaynaklar:

1. Editorial. Guidelines for reporting statistics in journals published by the American Physiological Society. *Adv. Physiol Educ* 28:85-87, 2004.
2. Kuzon WM, Urbanchek MG, McCabe S. The seven deadly sins of statistical analysis. *Ann Plast Surg*. 39(3): 265-272, 1996.
3. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: Current practice and problems. *Statist Med*. 21: 2917-2930, 2002.
4. Frasure SN, Lesperance F, Prince RH, Verrier P, Garber R, Juneau M, Wolfson C, Bourassa M. Randomised trial of home-based psychological nursing intervention for patients recovering from myocardial infarction. *Lancet*. 350:473-479, 1997.