ROMATOLOJİDE KLİNİK ARAŞTIRMALAR VE NESNEL TIP

EDİTÖR
HASAN YAZICI

BANU ÇAKIR
HANER DİRESKENELİ
AHMET GÜL
VEDAT HAMURYUDAN

ROMATOLOJİ ARAŞTIRMA EĞİTİM DERNEĞİ
RAED

ROMATOLOJİDE KLİNİK ARAŞTIRMALAR VE NESNEL TIP

© AresKare
Büyükdere Caddesi No: 15/A Hür Han B Blok Kat: 3 34387 Şişli - İstanbul Tel: 0212 343 82 30

Nisan 2007, İstanbul

Baskı ve Cilt: Kazmaz Matbaacılık
Eski Büyükdere Caddesi Cem Sultan Sokak 2/A 34416 4.Levent - İstanbul Tel: 0212 281 70 03-04

ISBN 978-9944-214-00-1

İÇİNDEKİLER

Romatolojide Klinik Araştırmalar ve Nesnel Tıp

Önsöz 5

Açılış konuşması 7

Geçerlilik, Tip I ve Tip II Hata ve Karıştırıcı Faktörler 9
Doç. Dr. Banu Çakır

Bir Demet Yanlılık 31
Prof. Dr. Hasan Yazıcı

Epidemiyolojik Çalışma Yöntemleri 37
Dr. M. Tevfik Dorak

Meta Analiz 41
Prof. Dr. John P. A. Ioannidis

Genetik Çalışmalarda Yöntem 45
Prof. Dr. Ahmet Gül

Genetik İlişki Çalışmalarında Karşılaşılan Sorunlar 55
Dr. M. Tevfik Dorak

İlaç Çalışmalarında Etki 59
Prof. Dr. Vedat Hamuryudan

İlaç Çalışmalarında Zarar 65
Prof. Dr. John P. A. Ioannidis

İyi Klinik Uygulamalar 71
Dr. Cem Özesen

Etik Komite Görüşü 77
Prof. Dr. Haner Direskeneli

Örnek İlaç Çalışması Sunumu: Gut Tedavisinde E Vitamini 87
Kısa Rapor

İndeks 93

ÖNSÖZ

Hasan Yazıcı

Hekimliğin nesnelleşmesi tüm dünyada görece yeni. Sayılar ve saymak ise nesnel, eleştirici ve araştırıcı hekimliğin vazgeçilemez öğeleri. Gerek tanı gerekse de tedavi yöntemlerinde son yıllarda adeta baş döndürücü bir hızla artan teknolojik ve moleküler biyolojik gelişmeler ise nesnellik gereksinimimizi, çoğu kez sanıldığı üzere, azaltmayıp aksine çoğaltıyor.

Araştırma bir yana, günlük hekimlik uygulamasında da gelişmekte olan ülke hekimi için ise nesnellik yaşamsal. Modern tıbbın kuşkusuz çok gelişmiş ama o oranda da pahalı tanı ve tedavi yöntemlerinden sağlık giderleri açısından en akıllı ve hastalarımız bakımından da en etkili yararlanmanın temelinde kuşkusuz nesnelllik var.

Nesnel tıbba ülkemizde gerekli önemin verildiği kanısında değiliz. 15-16 Nisan 2006'da yapılmış olan "I. Bodrum Klinik Araştırmalar ve Nesnel Tıp Günleri" bu boşluğu doldurma yönünde bir çabaydı. "Günler"in gördüğü ilgi bize onun bir çıktısı olan bu kitapçığı derleme cesareti verdi. "Günler"e ve kitapçığa katkısı olan herkese en içten teşekkürler.

Düzenleme Kurulu adına
Hasan Yazıcı

AÇILIŞ KONUŞMASI

Hasan Yazıcı

Bu oturumda planımız öncelikle epidemiyolojinin temel kavramları ve klinik çalışmalardan bahsetmek olacaktır. Ardından panellerimiz ve interaktif tartışmalı bir toplantımız olacaktır. İnteraktiviteyi sağlamak amacıyla olabildiğince dinleyici katkıları gözetilecektir. Düzeltmeler ve yorumlar araya girilerek yapılabilir.
Birinci konuşmacımız Banu Çakır bize temel kavramlardan bahsedecek.

GEÇERLİLİK, TİP I VE TİP II HATA VE KARIŞTIRICI FAKTÖRLER

Banu Çakır

Doçent Doktor, Hacettepe Üniversitesi, Tıp Fakültesi, Halk Sağlığı Anabilim Dalı

6 yıl önce Türkiye'ye geri döndüğümde, 1970'li yıllardan beri Halk Sağlığı Anabilim Dalları tarafından mezuniyet öncesi ve sonrası epidemiyoloji eğitimi formel olarak veriliyor olsa da, klinik hekimlerle işbirliğiyle çok fazla çalışma yapılmadığını gördüm. Bu üzüntü verici bir durum, çünkü meslek disiplinimizi yeterince iyi yayamadığımızı düşündürüyor. Öte yandan, tümüyle klinikte çalışan bir meslek grubunun bir buçuk gününü sadece epidemiyolojiye ayırıp ilgilenmesi, epidemiyoloji disiplininden biri olarak gerçekten çok sevindirici.

Klinisyen hekimler, aslında epidemiyolojiyi günlük işleri içinde, sürekli olarak kullanıyor. Belki adını koymadan, diğer bir deyişle "bu yöntemi, şu amaçla kullanıyorum" diyerek değil, ancak rutin işlerinin bir parçası olarak, araştırmalarını planlayıp uygularken, yazarken ya da hastaları ile ilgili makaleleri okurken sıkça kullanıyorlar. O halde, bu disiplini daha iyi tanımakta yarar var.

Ümit ediyorum ki, epidemiyolojiye olan ilgi tüm klinik dallara aynı şekilde yayılır ve klinik karar verme sürecinizde, kanıt elde etme ya da kanıtı yorumlama konusunda katkı sağlar. Epidemiyelojinin bir departmanın elinde olması yerine, tıpkı biyoistatistik biliminin kullanımında olduğu gibi, her departmanın kendi epidemiyologunu yetiştirmesi ve onun danışmanlığında çalışması bence çok önemli bir noktadır. Çünkü bu toplantıda uzun süre konuşacağımız şeylerden biri de, epidemiyolojik metodların hiçbirinin, örneğin, ne p değeri, ne güven aralığının, tek başına bir şey ifade etmiyor olmasıdır. Tıpkı tam kan sayımı, bir antikor değeri veya diğer laboratuar ölçümlerinin, klinisyenlere karar verme sürecinde yol gösterecek farklı göstergelerden yalnızca biri olması gibi. Günümüzde şöyle bir durum söz konusu: iyi medikal dergilerde bile, p değerleri ya da güven aralıklarının üzerinden gidip bir ilacın iyi olduğu ya da bir yöntemin üstün olduğu kararına ulaşılıyor. Oysa bu çok yanlış. Klinisyen hasta hakkındaki klinik karar hakkı ve hakimiyetini sadece bu sayılara dayandırırsa bilim gerçekleşemeyecektir, üstelik, çok yanlış yerlere de gidebilir. Klinik karar verme sürecinde hekim, öncelikle saptadığı durumun "klinik anlamlılığı"na bakmalı, daha sonra bu anlamlılığın "biyolojik" olarak açıklanabilirliğini değerlendirmeli ve son olarak, bu farkın istatistik açısından da

- Fibromiyalji, depresyon ve tiroid otoimmünitesi arasında bir ilişki var mı? *(Riberio RS, Proletti JA, 2004)*
- Sigara içmenin SLE'de hastalık aktivitesi ve kümülatif organ hasarı üzerindeki etkisi nedir?*(Ghaussy NO, et al., 2003)*
- COX2 spesifik inhibitörleriyle tedaviye başlayacak hastalarda yaşam boyu advers ilaç reaksiyonlarının derecesi, hizmet kullanımı ve hastalığın şiddeti nedir? *(Wolfe F, et al., 2002)*
- Yaygın ağrı ve Sjogren sendromu arasındaki ilişki nedir? Fibromiyalji ile bir bağlantı mevcut mu? *(Pendaris PT et al., 2001)*

değer taşıyıp taşımadığını incelemeli ve kararını tüm bu bilgiler ışığında vermelidir. Epidemiyoloji bilgisi, bu sıralamada, kanıt elde ederken kullanılan tasarlama, uygulama, analiz ve yorum süreçlerinin gerektiği şekilde olup olmadığını değerlendirebilmeniz için gerekir.

Önce geçerlilik, daha sonra tip 1 ve tip 2 hata, *power* ve en son olarak da karıştırıcı faktörlerden bahsetmek istiyorum. İster istemez İngilizce terimler kullanıyoruz, çünkü epidemiyoloji terminolojisi henüz Türkçe'de tam oturmadı. Farklı disiplinler bir yana, aynı disiplinden olanlar bile aynı kavrama başka bir isim verebiliyor. Terimlerin Türkçe karşılıklarını kullanmamız, böylece yaygınlaştırmamız aynı dili konuşmak açısından önemlidir. Türkçe kullandığım terim yadırgansa da, uygun olmadığı düşünülse de kaydedilmelidir; zaman içinde, daha uygun bir terminoloji sözlüğü yaratılabilir.

Araştırmacılar, bir çalışmaya başlarken önlerinde boş bir tünel ve içinde bazı sorular bulurlar. Bu sorulara örnek olması için, PubMed'de rasgele bir tarama yaptım ve romatoloji alanında yakın dönemde yayınlanmış bazı makaleleri buraya getirmek istedim. Burada geçen her örnek, bir romatologun çalışabileceği araştırma konularıdır. Örneğin, bunlardan ilki olan ve fibromiyalji, depresyon ve tiroid otoimmünitesi arasında bir ilişki olup olmadığını ve bu ilişkinin nasıl olduğunu araştıran çalışmanın benzeri yayınları Türkiye'de de gördüm. Yazıda depresyondan bağımsız olarak fibromiyalji hastalarında tiroid antiotoimmünitesi olduğu öne sürülmüş ve çalışılan ilişki için karıştırıcı faktörlerin neler olabileceği tartışılmıştı. Benzer bir araştırma konusu olarak sunmak istediğim ikinci örnek, sigara içmenin lupusta hastalık aktivitesine ve kümülatif organ hasarına etkisi ile ilgilidir. Başka bir örnek, son dönemde gündemde sık yer alan bir konu: COX2 spesifik inhibitörü verilen hastaların profillerinin diğer ajanlar verilen hastaların profilinden farklı olması nedeniyle, epidemiyolojide *confounding* ya da *bias* dediğimiz faktörlerin kontrolünün önemini vurgulayan hoş bir yazıdan oluşuyor. Son olarak taramamda, Sjögren ile yaygın ağrı arasında bir ilişki olup olmadığını ya da bunun aslında fibromiyaljiye mi bağlı olduğunu ele alan bir yazı buldum. Bu çalışmalardan herhangi biri sizin de ilgi alanınız ya da araştırma konunuz olabilir.

Ayrıca, bir kısmını birazdan örnek olarak da kullanacağız.

Burada vurgulandığı gibi, farklı araştırmacılar farklı konularda çalışma yapmak isteyebilir ve farklı epidemiyolojik yöntemlerden yararlanabilirler. Ancak, değişmeyen amaç şudur: çalışmamızda ne olursa olsun, 'gerçeği' öğrenmek istiyoruz. O gerçek, tıpkı yıllarca çok makbul olan, desteklenen, sonra gündemden düşen, şimdi tekrar makbul olan ve hatta bebeklere ilk 6 ay tek başına verilmesi önerilen anne sütü örneğinde olduğu gibi, **"bugünün gerçekleri"**. Ne yazık ki, "gerçek budur" demek ve bundan yüzde yüz emin olmak çok zordur. Biz, medikal çalışmalarımızın tümünde, gerçeği sadece "tahmin etmeye" çalışıyoruz: Elimizdeki en doğru, en yansız, en tam verilerle, gerçeği bulmaya çalışıyoruz. Gerçeği bulmaya çalışırken iki majör hata yapabiliriz:

1) Gerçek bir yerlerde mevcuttur. Örneğin, fibromiyaljiyle otoimmün tiroidit arasında bir ilişki vardır, ancak biz bunu bulamayabiliriz. Bu durumda ya Tip II hata yapıyor olabiliriz ya da sistematik bir hata vardır (herhangi bir *bias*). İngilizce'deki *bias* kelimesinin karşılığı olarak ülkemizde yanlılık, yan tutma ve taraf tutma terimlerinin kullanıldığını görebilirsiniz, ama *bias* teknik bir kelime olarak da Türkçe araştırmalarda oldukça yaygın kullanılır durumda. Ben, yanlılık ya da yan tutma terimlerini kullanıyorum, çünkü, istemli ya da istemsiz olarak, bir tarafı (yanı) tutmuş oluyorsunuz. Diğer sistematik hata kaynağı da karıştırıcı faktörler olabilir. Benim konuşmamı takiben Dr. Yazıcı *bias* kavramını tartışacak.

2) Araştırmacı bir ilişki saptar, ama gerçekte o tür bir ilişki yoktur. Bu da araştırmalarda sık karşılaştığımız ikinci bir durumdur ve gerçekte olmayan bir ilişki den bahseder oluruz. Bunun en yaygın olarak bildiğimiz formu, Tip I hatadır. Tip I hata önemli bir hata kaynağıdır, çünkü olmayan bir duruma var denirse bunun bir devamı olabilir. Diğer bir deyişle, hastaya ek/ileri tetkikler yapılabilir, başka kurumlara sevk edilir, zaman, insan gücü, para, eziyet (maddi ve manevi anlamda) söz konusu olur ve kişiler gereksiz yere evhamlanırlar. Tip I hata önlenemez ama araştırmadaki Tip I hatanın boyutu (p değeri) saptanabilir, dolayısıyla, bu hatanın ne kadar önemli olduğu tartışılabilir.

YA GERÇEK DEĞİLSE?

- Rasgele Hata ("şans"ın rolü)
  - Tip I hata (alfa)
- Sistematik Hata BIAS:
  - Seleksiyonda Bias
  - Bilgide (Ölçüm) Bias (Yanlış Sınıflandırma)
- Karıştırıcı Faktörler

Araştırmalardaki gerçeği tahmin etme çabamızın başarısını değerlendirmede iki kavramdan daha bahsedeceğiz: Biri *validity* (geçerlilik), diğeri *precision* (keskinlik). Şimdi dilerseniz bu kavramları ayrı ayrı detaylandırarak, tartışalım.

Her zaman gerçeği ölçmekten söz ediyoruz. Gerçek nedir? Nerede, nasıl çalışılır? Popülasyon, araştırmamızda incelediğimiz durum ya da ilişki hakkında yorum yapacağımız topluluktur. Toplumun tümü olabileceği gibi, çok az kişi içeren ya da çok özel bir topluluk da olabilir. Örneğin, hakkında çıkarım yapmak istediğimiz kişiler yoğun bakım hastaları, bir hastanede yatanlar ya da bir ülkedeki tüm 15-49 yaş kadınlar olabilir. Epidemiyolojide, istatistikçilerin aksine *universe* (evren) kelimesini kullanmıyor olma sebebimiz, sadece istatistik temsiliyeti arzu etmiyor olmamız ve çalıştığımız grubun da içinde olduğu ütopik, geniş, genel bir topluluğa ait çıkarımlar yapmak arzumuz olmamasıdır. Epidemiyolojide kullanılan popülasyon terimi risk altındaki toplumu ifade eder ve incelediğimiz çalışma sonuçlarına göre bu toplumdaki sağlık durumu hakkında genelleme yapılması amaçlanır. Risk kavramı da klinikteki risk kavramından biraz farklıdır; klinikte hastalardan bahsettiğimiz için risk kelimesi, durumu daha da ağırlaştıran faktörlerdir. Epidemiyolojide ise, risk ilgilenilen hastalık ya da durumla karşılaşıp, o hastalık ya da durumdan etkilenmesi olası anlamında kullanılır ve her zaman negatif/kötü bir durumu ifade etmez. Bazı kişilerde çalışılan durumun görülmesi biyolojik olarak mümkün değildir (kadınlarda prostat kanseri olmaması gibi), ya da bağışık olma durumu söz konusudur (kızamık hastalığı için aşılı olmak gibi). Dolayısıyla, kişinin o hastalığa yakalanma riski yoktur. O zaman bu kişilerin risk altındaki toplum dışında bırakılması gerekir.

Türkçe medikal yazımda sık yapılan bir hata var: *variable* (değişken) yerine parametre kelimesinin kullanılması. Örneğin, araştırmada hastanın belirli parametrelerine bakıldığından söz ederiz. Oysa, doğru ifade, araştırmada farklı değişkenlerin incelendiğidir. Değişken, değişebilen değer demektir. Hemoglobin değişebilir, kan basıncı değişebilir ya da benim kan basıncım başka bir arkadaşımın kan basıncından farklı olabilir.

Değişken farklı değerler alabilen şeylerdir. Bunlar ne olabilir? Sağlıkta, biyolojik, sosyolojik,

fiziksel, ekonomik değişkenler olabilir. Çoğu zaman, bir ya da birden çok değişkenin bir durumunu öğrenmek isteriz: ortalama kan basıncı, belirli bir antikorun düzeyi gibi. Ya da nokta değerler ile ilgileniriz: oran, orantı ve hız karşılaştırmaları gibi. Rölatif riskler, *odd ratio* (tahmini rölatif risk) kavramlarını aşağıda detaylandıracağız.

Peki, bunu nasıl yapıyoruz? Bunu herkeste yapmak kolay değildir; öncelikle hakkında fikir edinmek istediğimiz topluluklar genel olarak çok kalabalıktır, dağınık yayılım gösterir ve hepsine ulaşmak mümkün değil ya da zordur, karşılanamayacak düzeyde zaman, para ve insan gücü gerektirir. Sık karşılaştığımız bir talep, herhangi bir hastalığın prevalansını bulmaktır. Örneğin, Çernobil olayından sonra Hopa'da ya da Karadeniz Bölgesi'ndeki insanların bu olaydan ne kadar etkilenmiş olduğunu bulmak ya da daha spesifik bir örnek olarak, bu bölgede tiroid kanseri prevalansını saptamak gibi. Bir hastalığın yaygınlığını veya görülme olasılığını belirlemek için kaç kişiye gitmek gerekir? Seksenbin, doksanbin, yüzbin haneye gitmek gerektiğini düşünün.

Örneğe çıkan kişi sayısı arttıkça gerçeği tahmin olasılığımız da artacaktır. Bu anlamda, örneklem büyüklüğünü artırmak olumlu sonuçlar verir. Ancak, örneklem sayısı büyüdükçe çalışmanın veri kalitesinden ödün veriyoruz; çünkü, belirli sürede veri toplamayı tamamlayabilmek için on ev olsa bir anketör, yüz ev için belki üç anketör, ama bin ev olunca onbeş anketörün eşzamanlı çalışması gerekecektir. Bu durumda verinin kalitesi değişecektir. Standart eğitim, ön deneme ve araştırma sırasında denetimlerle bu kaliteyi artırmak olasıdır ancak yüzde yüz tutarlılık mümkün değildir. Onun için hem istatistik, hem epidemiyoloji, hem de klinik açısından çoğu zaman toplulukların tüm üyelerine gitmek yerine, temsili örnek(lem) seçmekten yanayız.

Örnek seçiminde iki önemli özelliğe dikkat etmek gereklidir:

1) çalışmaya katılacak kişilerin yeterli sayıda olması ve
2) toplumu temsil edici olması.

Yani, toplumun alt gruplarının nelerden etkileneceğini düşünüyorsak o alt grupların örnekte yeterince temsil edilmesi gerekir.

DOĞRU TAHMİN
GEÇERLİLİK
- Sistematik hata olmaması
  validus güçlü (Latince)
  "bir çalışmadan çıkartım yapabilme derecesi (güçlü)
  bias
  karıştırıcı faktör

KESKİNLİK
- Rasgele hata olmaması
  "keskin tanımlanmış yada belirtilmiş olma özelliği"
  (Standart hata / sapma ile)
  tip 1 ve tip 2 hata

Çalışma grubundan elde ettiğimiz değerler ile (istatistik) toplumda karşılık gelen değerleri (parametre) tespit etmeye çalışırız. Bu durumda, gerçek değerlerin ne olabileceğini tahmin ederiz. Bu tahmini değer ne olabilir? Tek bir nokta değer olabilir, güven aralığı olabilir, *odds ratio* olabilir veya hız olabilir. Burada önemli bir soru sorabiliriz: Bunu ne kadar doğru ölçeceğiz ki, doğru yorumlar yapabilelim?

Elimizin altında iki kavram var: *validity* ve *precision*. John Last'ın, epidemiyoloji sözlüğünden bazı tanımlar kullanmak istiyorum. *Validity* (geçerlilik), çalışmada sistematik bir hatanın olmaması demektir. Latince güçlü sözcüğünden gelir. Oluşturduğunuz örnekten bir sonuca ulaşıp, bunu topluma yansıttığınızda yaptığınız varsayımın (çıkarımın) ya da *interference*'ın ne kadar doğru olduğunu gösterir. Bildirdiğimiz sonuçların gerek toplumun tümü için gerekse çalışılan grup bazında doğru olması esastır. Örneğin, fibromiyalji ile otoimmünite arasında ilişki olduğu ne kadar doğrudur? Hem kendi örneğim içinde hem toplum bazında doğru mu ölçülüyor? Doğruluğu tehdit eden sistematik hata kaynakları taraf tutma *(bias)* ve karıştırıcı faktörlerdir.

Precision ise rasgele hata ile ilgilidir; ölçümün gerçeğin çevresinde bir yerde dağınık değil de daha keskin olması istenir. Örneğin, hesapladığınız rölatif risk değerinin geniş değil de dar bir % 95 güven aralığı olması gibi. Diğer bir deyişle, "bu çalışmayı benzer şartlarda 100 kez tekrarlamış olsak 95'inde bulacağımız değerler birbirine yakın olurdu," sonucunu ifade etmeye çalışıyoruz. *Precision*'ı tanımlarken genel olarak, söylenen ya da ölçülen şeyin sınırlarının olabildiğince belirgin olmasından bahsediyoruz. Örneğin, kaç yaşında olduğunuz sorulduğunda yaşınızı tam olarak belirtmeniz: "25-30 arası" değil de "26 yaşındayım".

DOĞRU TAHMİN
KESKİNLİK
Rasgele hata olması
Tip 1 ve tip 2 hata
- Ölçüm ya da tahminde keskinlik rasgele hatanın azlığı YA DA tahmin edilmeyecek yanlışlıkların minimuma indirilmesine karşılık gelir
- Nasıl artırılabilir?
  Örneklem boyunu artırarak çalışmanın etkinliğini (birey başına maliyete karşı bilgi miktarını) artırarak

Precision rasgele hatanın en aza indirilmiş halidir. Bunun için iki şey yapılabilir. Birincisi, örnek büyüklüğünü artırmaktır: Örneğin, küçük bir örnek alındığını varsayalım, yalnız orta boylular ya da uzun boylular gibi. Oysa, ne kadar çok kişiye gidilirse toplum taraması o kadar belirgin olur. Her kesimden insan alındığı için elde edilen sonuç da daha keskin olur. Bir de çalışmanın *efficiency*'sini arttıracak bazı şeyler yapılabilir. *Efficiency* (verim), kabaca, birim kişiden bilgi

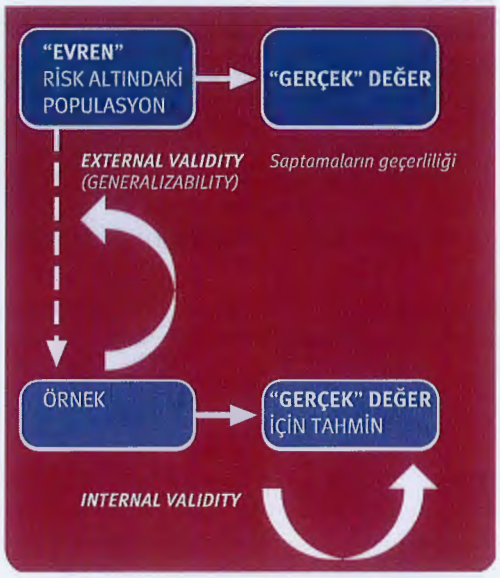
alındığında, kişi başına alınan bilginin, harcanan emeğe ne kadar karşılık geldiğini belirtir. Toplumdan tabakalı örneklem seçimi verimi artırmak için kullanılan yöntemlerden biridir.

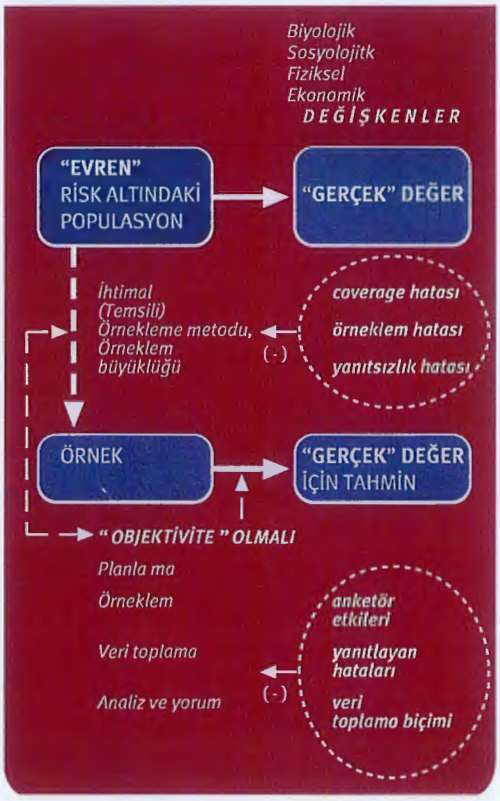
Validity (geçerlilik) ise, ikiye ayrılır: internal (iç) ve eksternal (dış) validite. Burada iki kavram karışabilir: Biri, çalışmanın, diğeri ise ölçümün geçerliliği. İngilizcede her ikisi de *validity*. Türkçe'de ölçümün *validity*'sini, yani yapılan ölçümün ne kadar doğru olduğunu ifade etmek için geçerlilik kelimesini kullanıyoruz. Bu anlamda geçerlilik, duyarlılık, seçicilik, pozitif prediktif değer ve negatif prediktif değer gibi kavramları da kapsıyor. Şu anda ise, çalışmanın geçerliliği, yani elde edilen sonuçlara göre yapılacak yorumların/çıkarımların (yani *interference*'ların) validitesinden bahsediyoruz.

Bir çalışmada toplumdan özel bir grup seçilirse, çalışma sonuçları o gruba özeldir, diğerleri hakkında iddiada bulunulamaz. Örneğin, sadece kadınlarda bir çalışma yapılıyorsa aynı grubun erkekleri için hiçbir şey söylemek mümkün değildir. Dolayısıyla, kim hakkında bilgi almak isteniyorsa ona göre örnek seçilmelidir. Özellikle, kesitsel tipteki prevalans araştırmaları bu nedenle toplumun tümünü temsil eder sayıda örnek seçimini gerektirir. Ama çalışma yaparken tek amaç genellemek değildir, kesitsel çalışmalar da buna dahildir. Bazı kesitsel çalışmalarda amacın prevalans tayini değil belirli etken-hastalık ilişkilerini araştırmak olduğu unutulmamalıdır.

İlk prevalans çalışması Fransa'da nüfus sayımı amacıyla yapılmıştır. Fransa'nın değişik coğrafi bölgelerinde 200 yıl kadar önce, seçilen bölgelerin tümüne giderek ve coğrafi bölgeleri şimdiki küme örneklemi gibi ele alarak sayım yapılmış ve bu sayımdan Fransa nüfusu bulunmuştur. 1930'lu yıllarda ekonomik ve politik olarak ya da görüş almak için bazı çalışmalar yürütülmüştür. Bunların tümü prevelans araştırmalarıdır ve kesitsel tarzda yapılmıştır. Önceleri kapı kapı dolaşıp sağlık araştırması yapılarak hasta olan kişi sayısına ulaşılırken, daha sonra bunun çok masraflı olduğu görülmüş ve etken sonuca bakmak, bazı şeyleri ilişkilendirmek gibi farklı bir yöntem geliştirme ihtiyacı doğmuştur. Şöyle bir yaklaşım doğru değildir: "bu toplumun yüzde yüzüne gidilmediği için çalışma genellenebilir değil, dolayısıyla değersiz". Eğer grup içinde bazı

DOĞRU TAHMİN
GEÇERLİLİK
Sistematik hata olmaması
Bias
Karıştırıcı faktörler
Çıkartımların geçerliliği
INTERNAL VALIDITY
Çalışma popülasyonunun üyeleri (kaynak)
EXTERNAL VALIDITY
Popülasyon dışı unsurlar (hedef / süper-popülasyon)





etken-sonuç ilişkilerini irdelemek ve bu etken-sonuç ilişkisinin büyüklüğü ile ilişkili faktörleri öğrenmek istersek (iç geçerlilik, *internal validity)*, mutlaka toplumu temsil etmek gerekmeyebilir. Bir çalışmanın iç geçerliliği, veri toplama kalitesi, taraf tutma ve karıştırıcı faktörlerin uygun kontrolü vb. faktörler ile ilişkilidir ve çalışma kalitesi bunların sağlandığı ölçüde artacaktır. Öte yandan, eğer topluma ait bir prevalans değeri verilecekse, temsili bir örnek üzerinde çalışılmalıdır; çalışmanın dış geçerliliğinin *(external validity)* yüksek olması için sayının da yeterli olması gerekir.

Özellikle lojistik regresyon ve diğer çok değişkenli modelleme yöntemleri geliştikten sonra, bir kohort grubunun içinden bir kesit alıp orada bazı analizler yapılması mümkün olmuştur. Bu çalışmanın genellenebilirliğini düşürse de iç tutarlılığını bozmaz.

Internal validity, bir çalışmanın belirli bir iç grubunda çalışılan, üzerinde birebir ölçüm yapılan insanlar bazında söylenen ilişkilerin geçerliliğidir. Çalışmalarda öncelikle iç geçerliliğin sağlanıp sağlanmadığına bakılmalıdır. Bir sonraki adımda tartışılması gereken, *external validity*, yani elde edilen sonuçların toplumu ne denli yansıttığıdır.

Burada toplum tanımında bir sıkıntımız var. Ben geçmiş deneyimlerimde hiçbir zaman toplumun %80'inin üstüne ulaşamadım. Klinik araştırmalarda grubun bir kısmına ait verinin eksik olmasında bence doktorların da payı var çünkü hastanın kayıtları kaybolmasın diye bazı dosyalar masalarda tutuluyor. Ancak, tek neden bu değil; bir çalışma yaptığınızda bir kişi yanıt vermek istemeyebilir ya da o kişiye gerçekten ulaşamayabilirsiniz. Burada *coverage error* (kapsayıcılık hatası) dediğimiz kavram devreye giriyor: o toplumun örnek seçilen altgrubu, toplumu temsil etmiyor olabilir. Örneğin, sadece telefonu olanlarla anket yapıyorsanız o toplumda telefonsuz olanlar çalışma dışında kalır, kapsayıcılıkta bozukluk vardır, yani toplumun tümüne ulaşılamaz.

Bir de *sampling error* vardır: örnek seçtiğiniz altgruplar toplumun tümünü iyi yansımayabilir. Bunların hepsi çalışmanın toplumun tümünü temsil etmesine engeldir, çünkü toplumun bir kısmı hakkında bilgi alınamamaktadır.

Bu verisi eksik kişiler, bizim çalıştığımız etkenle sonuç ilişkisi açısından çalışmaya katılmış kişilerden farklı özellikler taşıdığı nisbette, çalışma sonuçları toplumu temsil edemeyecektir. Bunun dışında, anketörün ya da bilgi toplayan kişinin etkileri ve cevap veren kişiyle ilgili faktörler de vardır. Cevap veren kişi yaklaşımı, soruları soruş tarzını ya da cevaplara gösterilen tepkileri beğenmeyebilir ya da, özel bir konu (örneğin, tüberküloz, sosyal bir hastalık) söz konusu olduğunda olduğu için sorularınıza yanıt vermeyebilir. Bu durumlarda cevap almanın özel yöntemleri vardır ama her zaman bazı bilgileri alamamak mümkündür. Bunlar da çalışmaları etkiler.

Tartışılması gereken bir konu da dış *validity* ve genelleştirilebilirlik *(generalizability)* kavramlarının farkıdır. Örnek seçimi ve örneğin nasıl seçildiği, neleri etkilediği konusunda epidemiyologlarla istatistikçiler arasında iki farklı görüş vardır; bir grup *design-based approach*'u savunurken diğer grup *model-based approach* taraftarıdır.

Design-based approach yıllardır Türkiye'deki çalışmalarda kendini göstermektedir. Burada ilgi duyulan popülasyondan örnek seçilir ve sonuçlar örneğin seçildiği toplumdaki kişilere genellenir. Örneğin, "Bu toplumda hastalık şu faktörler ile ilişkilidir," denir.

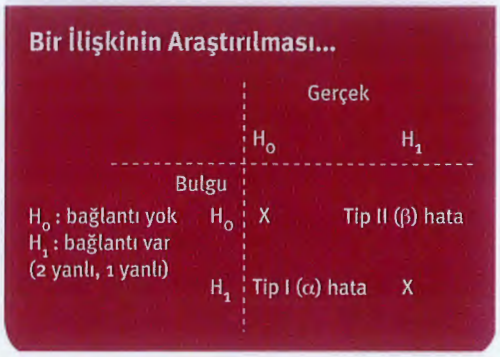
Model-based yaklaşımı savunanlar ise, bu çıkarımların toplumun genelini ifade etmediğini, sadece cevap verenlere özel olduğunu savunur; ancak, çalışma grubu da sonuçta benzeri özelliklere sahiptir, daha büyük, sınırlarını belirlemediğimiz, ancak var olduğunu kabul ettiğimiz bir topluluğun alt grubudur ve amaç çalışma grubundan elde edilecek sonuçların ışığında bu büyük evren hakkında fikir söyleyebilmektir. Bu varsayımı güçlendirmek için, toplumlardaki hastalıklar ve ilişkili olduğu faktörler hakkında karar süreçlerinde sadece kendi toplumumuzdaki çalışmalara bakamadığımız iddia edilir; çoğu zaman bu tür özel çalışmalar mevcut bile değildir. Hatta bazen hayvan deneyleri sonuçlarına göre varsayımlarda bulunulur. Yani, farklı çalışmalara ait sonuçlardan yola çıkarak klinik karara vamak mümkündür.

John M. Last A Dictionary of Epidemiology, 4th ed., 2001

- Hata bir çalışma ya da deney sırasında yanlış veya hatalı olarak elde edilen sonuçtur. Birkaç tipi vardır.
- Rasgele hata herhangi bir ölçüme veya değişkene,"şansa bağlı" olarak belirgin bir bağlantısı olmayan farklılık bölümüdür.
- Sistematik hata genelde farkedilebilir bir kaynağa sahiptir (örn, yanlış bir ölçüm). Bir yönde sürekli olarak hata oluşturur.

GEÇERLİLİĞİ TEHDİT EDEN FAKTÖRLER

- Rasgele hata (Şans)
- Karıştırıcı
- Seçimle ilgili "Bias"
- Bilgilendirme ile ilgili "Bias"



O halde bir örneklemden, ya da üstünde birebir ölçüm yapılan kişilerin değerlerinden yola çıkarak bir istatistiksel model geliştirilebilir ve bu model ile bulunan sonuçlar açıklanabilir. Örneğin, bu model koroner kalp hastalığının belirleyicileri arasında kolesterolün değerini ifade edebilir. Bu şekilde, bazı temel özelliklerini bildiğimiz durumlarda, farklı topluluklarda da aynı modelden yola çıkarak, kolesterol-koroner arter hastalık ilişkileri hakkında fikir üretilebilir. Bu tür *model-based* yaklaşımlar için ciddi istatistik desteği gerekir, paket programlar yeterli olmayabilir.

Onun dışında en büyük sıkıntı, yapılan modelin varsayımlarının yanlış kurulması durumunda sonucun tamamen yanlış olabilmesidir. Şunu unutmamalıyız: ilişkilendirme değil de prevalansa bakılıyorsa gerçekten toplumu temsil eden bir örnek seçilmeli ve bu topluma yansıtılmalıdır. Öte yandan, hangi yaklaşımı savunursak savunalım toplumdan temsili bir örnek üzerinde çalışılması zorunludur. Toplum için bir şey söylemek için değilse bile modelin içindeki değişkenleri doğru kurmak için bu şarttır. Son olarak, *design-based* yaklaşımı savunanlar çalışma sonuçlarından yorum yaparken *external validity*, *model-based* yaklaşımı savunanlar ise daha çok *generalizability* terimlerini kullanmaktadır.

Last'ın kitapçığından aldığım tanımları ile şu üç kavramı tartışmak isterim: Hata, rasgele hata ve sistematik hata. Rasgele hata bir kişinin bir ölçümü yaparken herhangi bir şeye bağlı olmaksızın, özel bir nedeni olmaksızın ya da çalışılan etken-sonuç ilişkisini etkilemeksizin, daha çok şansa bağlı olarak varyasyon göstermesidir. Bu engellenemez ama boyutuna karar verilebilir (p değeri); bu boyut istenenden (alfa değeri) fazlaysa reddedilir, az ise istatistik açısından anlamlılıktan söz edilir.

Sistematik hata ise daha farklı bir hata tipidir; sistematik olarak yanlış ölçüm yapılır ve hatalı olarak var olan ilişkiler yok, yok olan ilişkiler var, var olan ilişkiler az ya da çok gösterilebilir.

Sağlık araştırmalarında yapılan hatalar rasgele ya da sistematik olabilir. Rasgele hata keskinlik ile ilgili, sistematik hatalar ise geçerlilik ile ilgilidir. Bir çalışma yaparken bir yokluk hipotezi ya da *null hypothesis* vardır.

Yokluk hipotezinde çalışılan etken ile sonuç arasında bir ilişki olmadığı iddia edilir. Örneğin, Sjögren hastalığında, diğer hastalıklardan bağımsız olarak yaygın ağrı olmadığı iddia edilebilir. Gerçekten de yoksa ve biz de yok bulursak sorun yoktur. Ancak gerçekte böyle bir ilişki var ve biz de bulursak yine sorun yoktur. Buradaki iki sorun var olan bir ilişkiye yok demek ve yok olan bir ilişkiye var demektir.

Önce anlaşılması daha kolay olanını inceleyelim: Var olan bir ilişkiye yok demenin bedeli nedir? İlişkiler bulunamaz, başka araştırmacıların tekrar çalışması gerekir ya da unutulur. Çalışma bu anlamda en çok örnek büyüklüğünden etkilenir; onun için örnek büyüklüğü seçerken ne kadarlık bir güç *(power)* istendiğine karar verilmelidir.

Power çalışmanın gücüdür: 1β yani, tip II hatanın 1'den çıkarılmış halidir. Eğer yüzde yirmi hataya izin verilirse, ki maksimum bu kadara izin verilir, çalışmanın var olan bir ilişkiyi gösterme gücü % 80 demektir. Örneğin, Sjögren'de gerçekten yaygın ağrı varsa bunu çalışmada % 80 olasılıkla bulmak için gerekli düzenleme yapılmalıdır. Peki, seçilecek örnek büyüklüğü ne olmalıdır? Güç, örnek büyüklüğünden, hastalığın toplumdaki sıklığından, ilişkinin büyüklüğünden, çalışmanın tipinden, ne kadar kesin ölçüm yapmak istediğinizden etkilenir.

TİP I (ALFA) HATA

- Gerçek null bir hipotezin inkar edilebilmesi (oranı)
- Olmadığı halde bir farkın olduğunun açıklanması
- Şansa veya örneklem değişkenliğine bağlı olarak bir çalışmadan yanlış bir pozitif "outcome" çıkması
- Anlamlılık düzeyi (Fisher)-çalışmadan önce araştırmacı tarafından belirlenen
- Bir "null" dağılım tarafından temsil edilen, gelecekteki bir dizi "outcome'" olasılığı

BİR ÇALIŞMANIN GÜCÜ

"Varsa bir ilişkinin gösterilebilmesi"
güç= 1- β (tip II hata)
Şunlara bağlıdır

- Örneklem büyüklüğü
- Araştırılan tablonun sıklığı
- Etkinin boyutu
- Çalışma tasarımı
- İstenen keskinlik
- İzin verilen Tip I hata

P DEĞERİNDEN NEDEN FARKLIDIR?

~~"otomatik" p değeri kullanımı~~

- Veriler elde edildiğinde, "null" hipoteze karşı kanıtın ölçümü olarak p değeri kul lanılır
- P > α ise sonuç anlamlıdır, örn., null hipotez'in "gerçek olması olası değildir"
- P değeri esnek bir "inferential" değerdir
- Birçok durumda günümüzde kullanılan hipotezelerle uyumsuzdur

P DEĞERİ

- R. A. Fisher, "modern istatistiğin babası"
- Gözlemlerden objektif, kantitatif bir metodla, süreçleri çıkartım yapıcı sonuçlara yönlendirmek istemiştir.
- Bayes formülüne (önceki olasılığın, sonra ki olasılığa konversiyonu) inanmamıştır
- P değerini ilk kullanandır

FISHER'İN P DEĞERİ

- P değeri tek bir çalışmada, veri ışığında "null" hipotezin kredibilitesini yansıtmak için kullanılan bir kanıt ölçümüdür.
- Kanıtın ölçümü olarak oldukça esnektir.
- Null'i hiçbir zaman kabul etmez, sadece reddedemez

Tip I (α, alfa) hatanın farklı kaynaklarda farklı ifadelerle tanımlandığı görülebilir. Tip I hata, gerçekte olmayan bir ilişkiye var demek, gerçekte gruplar arasında fark yokken fark var demek veya örneklemdeki değişiklikler ya da şans eseri çalışma sonucuna yalancı pozitif sonuç almak demektir. Fisher, alfa hatayı anlamlılık düzeyi olarak tanımlamıştır: α hata, çalışmadan önce çalışmacı tarafından belirlenen bir düzeydir ve normalde olması gereken bir dağılımın iki tarafında kalan kısmı ifade eder.

P değeri sıklıkla hatalı kullanılır. P değerini bulabilmek için veri setinin elinizde olması gerekir. Alfa değeri gibi baştan karar verilemez; çalışmada elde edilen veriler ışığında bir p değeri hesaplanır ve sonucunda yokluk hipotezinin aleyhine veya lehine kullanılır ya da kullanılmaz.

Eğer bir çalışmadan elde edilen p değeri α değerinden küçükse, yokluk hipotezi reddedilir ve sonucun istatistik açısından anlamlı olduğu söylenir. Dolayısıyla, bulunan ilişkinin gerçekte var olduğu kararına varılır.

Fisher rasyonel olmak, ölçüm değeri yapmaktan söz eder. Buradaki önemli özellik şudur: p değeri tek bir çalışmadan elde edilir, ölçülmüş bir değerdir ve yokluk hipotezinin lehine kullanmak için yapılır. Fisher çok esnek kullanılmasından, klinisyenin mantıklı olanın % 5 mi yoksa % 8 mi olduğuna kendisinin karar vermesinden yanadır. P değeri sadece yokluk hipotezini reddedebilir. Yokluk hipotezinin reddi ise herhangi bir anlama karşılık gelmez. Başka bir deyişle, alternatif hipotez yoktur.

Neyman ve Pearson ise, araştırmalarda yokluk hipotezine karşı alternatif bir hipotez olmasını ve bunu da sınamaya yönelik bir test geliştirmeyi hedeflemişlerdir. Ancak Neyman bir kritik bölge olduğundan söz eder. Buna göre, yokluk hipotezi ile alternatif hipotezin bir yerde örtüşmesi gerekir ve araştırmada saptanan p değeri belirli bir kesme noktasının dışındaysa yokluk hipotezi reddedilir ve alternatif hipotez lehine karar verilir.

Ancak, Neyman ve Pearson'ın kavramlarında da şöyle bir sorun vardır: belirli bir bölgedeyse hipotez kabul edilmez ama bunun da sınırı yoktur. Yani, orijinalinde, Fisher'in p değeri gibi, "yokluk hipotezi için hesaplanan p değeri 0.3'tür" gibi bir hesap söz konusu değildir.

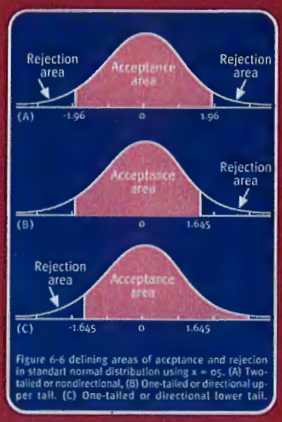


Figure 6-6 defining areas of acceptance and rejecion in standart normal distribution using α = 05. (A) Two-tailed or nondirectional, (B) One-tailed or directional upper tail. (C) One-tailed or directional lower tail.

Günümüzdeki tanımı aynıdır ancak Fisher'in nasıl kullanılması ve ifade edilmesi ile ilgili fikirleri farklıdır

"p değeri "null" hipotez altında belirli bir deneysel gözlem olasılığına, artı daha ekstrem olanlara eşittir

NEYMAN-PEARSON HİPOTEZ TESTİ

- Bir değil, iki hata oranı; Z>1.96 gibi özet istatistik için "kritik alan"ı ifade eder.
- "Kritik alan"a düşen, null'u reddeden, alternatifi yada tam tersini kabul eden sonuçtur.
- Kanıtın bir ölçümü değildir; sonucun kritik alana düştüğü söylenebilir, nereye düştüğü değil

NEYMAN-PEARSON HİPOTEZ TESTİ TEST

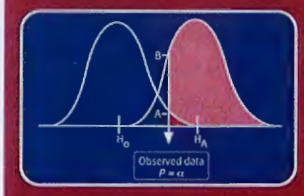
- .. Olasılık teorisine bağlı bir test tek başına gerçeğe veya hipotezin yanlışlığına dair değerli bir kanıtı sunamaz.
- Ancak testlerin amacına başka bir açıyla bakabiliriz. Ayrı ayrı her testin doğru ya da yanlış olmasını ummak yerine, onlarla ilişkili davranışımızı yönlendiren kuralları araştırabiliriz - bir tipin H'sinin reddedilip reddedilmeyeceği...

Kaynak: Neyman J, Pearson E. Philos Trans R Soc London A 1933;231:289-337.

NEYMAN-PEARSON HİPOTEZ TESTİ

- "Inductive" değil, "deductive" mantık
- "Inductive" davranış (mantık değil)
- Fisher hipotez teslerini "kabul prosedürleri" ve "karar fonksiyonları" olarak ele alır-sadece matematiksel değil

gözlemlenen hata oranı olarak p değeri...



Ayşe 100 öğrenci içinde 15. gelir:

- *İlk % 15'te midir?*
- *Bu eşik sınır bağımsız olarak mı konmuştur?*

Anlam düzeyi (tip I hata oranı - örneklem hatasına veya şansa bağlı olarak)
α, "null" dağılımın "kuyruk alanının" temsil ettiği bir dizi sonucun olasılığıdır - koşulsuz
P değeri, kuyruk alanının tam olarak gözlenen olasılığı kısmında yer alır. Koşullu

BİR HİPOTEZİN OLASILIĞI

- Doğru olma olasılığı DEĞİLDİR
- Olasılık hipotezin veriyle desteklenip desteklenmediği ve ne kadar desteklendiğidir.
- "Gözlemlenen veriler ışığında, olasılık gözlemlenen verilerin ihtimaliyle orantılıdır"
- Fisher'e göre, "olasılık değerlendirilen olasılıklar arasında doğal bir tercih sırası oluşturur"

BİR HİPOTEZİN OLASILIĞI

- Olasılıkların oranı (ya da logaritma) iki hipoteze verilerin sağladığı göreceli kanıta dayalı desteğin ölçüsü olarak kullanılır.
- P değeri < 0.05 ise: "sıfır hipotezinde ihtimali %5 ya da daha düşük olan bir olay meydana gelmiştir"

P DEĞERİNİ BİLDİRİRKEN KARIŞIKLIK...

- P > yada > 0.05, *(saf hipotez testi perspektifi, doğrudan a ile bağlantılı)*
- Tam p değeri söylenir, örn., p = 0.03 *(bağlantıyı koparır)*
- "Alfa tarzı" p < 0.05, p < 0.01, p > 0.001
- P değeri anlamlılığı *, **, *** olarak gösterir
- P < 0.03 (tam p değerini eşitsizlik olarak ifade ediyor)

P DEĞERİNİN KULLANIMI-SONUÇ

- P değeri şansın rolünü belirlemede katı bir kural değildir, ama şansın bulguları açıklayabileceğini ifade eder.
- Bir bağlantının istatistik açısından anlamlığı biyolojik ya da klinik anlamlılığından ayırt edilmelidir.
- P değeri bildirilecekse hesaplanan tam değer olarak bildirilmelidir, örn., p=0.023
- Ama, en iyisi güvenlik aralığı kullanmaktır...

Zaman içinde p değeriyle hipotez testi bir araya gelmiş ve genellikle, günümüzde de olduğu gibi, p değeri α yerine kullanılır olmuştur.

Bazı klasik ders kitaplarında bile p değeri tam olarak Tip I hatayı tarif eder. Oysa α çalışmaya başlamadan verilen bir karar değeridir; p ise verilerin elde edilmesinden sonra hesaplanır. Alfa *unconditional*, P değeri ise *conditional* data'yı verir.

Son olarak, verilerin ışığında öne sürdüğümüz hipotezin doğru olma yüzdesi *likelihood* değeri ile hesaplanabilir. Biyometri ve genetikte çalışmış bir istatistikçi olan Fisher elindeki bazı gözlemlerden yola çıkıp yaşamla ilgili objektif bir şeyler söylerken güvenilir olmak istemiş. Burada 'Bayes Teoremi' devreye giriyor. Bayes Teoremi'nde sonucu söylemek için önce o olasılığın, o hastalığın ya da durumun zihninizde olması ve çalışmayı yapınca düzeltilmesi gerekir. Fisher'in geliştirdiği matematiksel olabilirlik (*mathematical likelihood*) kavramı, eldeki veri seti ile değişik düzenlemeler/modeller yaparak elde edilen sonuçları en iyi açıklayacak modeli bulmaya dayanır. Buna göre, eldeki verilere göre sonuca ulaşma olabilirliği kestirilir; p < 0.05 bulunursa, yokluk hipotezi doğru kabul edilirse, şans eseri ortaya çıkması %5 olan bir durumdan söz ediliyor demektir.

Pratikte p değerinin çok farklı tarzda ve çoğu durumda hatalı kullanıldığını görüyoruz. Örneğin, p değerinin alfa'dan (örneğin, % 5'ten) küçük olduğunu belirtmek yeterli değildir; % 1 mi yoksa % 0.001 mi olduğu da belirtilmelidir. Klinisyen kararını bu bilgiler ışığında kendi verebilmelidir.

Şöyle bir örnek verilebilir: okulunuzdan transkript istediğinizde kaçıncı olduğunuzu sorabilirsiniz. Yüz kişinin içinde onbeşinci olduğunuzu varsayalım. Bunu bildirirken en üst % 15'de olduğunuzu söylemek doğru bir ifade değildir, çünkü o zaman en üst % 15'in neresinde olduğunuzu belirtmek gerekir. En üst % 15 yerine onbeşinci demek gerekir. P değeri de böyle bir kavramdır, yani p değerini 0.03 hesapladıktan sonra, %5'ten küçük olduğunu belirtmek yerine tam olarak değerini vermek gerekir. *Roving alfa* denen ve yıldızlarla gösterilen bir ifade tarzı da vardır: örn., ** ise p % 5'in, *** ise p % 1'in altındadır.

Başka bir hatalı gösterim de, % 3 bulunmuş bir değeri ifade ederken p < 0.03 olarak kullanmaktır.

Epidemiyolojide en doğru yaklaşım, p değeri mutlaka verilecek ise, hesaplandığı değer ile bildirilmesi, tercihen p değerleri yerine güven aralıklarının (genelde % 95) kullanılmasıdır. Zira, nokta değerler güven aralıkları ile birlikte kullanıldığında hem istatistik açısından anlamlılığı değerlendirebilir hem de bu anlamlılığın ne aralıkta değiştiği hakkında bilgi edinilebilir. Öte yandan, tıpkı p değeri gibi örnek büyüklüğünden de etkilenir.

Çalışmaya katılan sayısı (örneklem) ne kadar büyükse p değeri o kadar anlamlı hale gelecektir. Hatta böyle durumlarda p değeri bir ilişkinin olduğundan 34 kat daha fazla anlam kazanmasına neden olabilir, örneğin, 0.05 bulunduğunda, aslında *likelihood*'la bunun karşılığına bakılırsa 0.20'ye, 0.15'e denk olduğu bulunabilir. Bu nedenle, özellikle lojistik regresyon yaparken önce ikili ilişkilere bakılır, onların içinden p < % 20'lere kadar olanları anlamlı olabilir diye modele alınır. Başka bir deyişle, p değerleri biraz *overestimation* yaratır, Tip I hatayı güçlendirir.

Güven aralığı değerleri kullanmanın avantajı, istatistik açısından anlamlılık ile hesaplanan nokta değerin (belirli bir güven aralığı içinde) alabileceği minimum ve maksimum değerleri de vermesidir.

Bir yazıda "istatistik analizi fibromiyalji ile tiroid otoimmünitesi arasında bir bağlantı olduğunu gösterdi (OR=3.8)" ifadesi olduğunu ve bu çalışmada p değerinin 0.03 olarak bildirildiğini varsayalım. Bunun anlamı şudur: fibromiyaljiyle tiroid otoimmünitesi arasında bir ilişki yoksa, yani yokluk hipotezi doğruysa, çalışmada OR=3.8 ve üstündeki bir ilişkiyi şans eseri bulma olasılığı 0.03'tür. Klinisyen olarak 0.05'i anlamlı sayıyorsak, bunun şans eseri olmadığına karar verir ve bu bulguyu istatistik açısından anlamlı kabul ederim.

Güven aralığı değeri kullansak ne olacaktı? Varsayalım ki, çalışmada bildirildiğine göre, "antitiroid peroksidaz ya da antitiroglobülin antikorları olan kişilerde fibromiyalji görülme rölatif riski, olmayanlara göre 3.8 kat daha fazladır ve %95 güven aralığıyla bu değer 1.54 ile 10.13

GÜVEN ARALIĞI

- *"Gerçek değer"in yattığı aralığı, belirli bir güven ya da ihtimal ile sunar*
- *İstatik açısından anlamlılık (şansın rolü) hakkında, ama aynı zamanda, etkinin boyutunun aralığı hakkında da bilgi sağlar*

GÜVEN ARALIĞI

- *Röllatif Risk için % 95 GA:*
  $(RR).\ exp\,[\pm 1.96\sqrt{Varyans(lntRR)}]$
- *Odds Ratio için % 95 GA:*
  $(OR).\ exp\,[\pm 1.96\sqrt{Varyans(lnOR)}]$

Unutmayın:
GA'nın eni veri toplama süreci, seçilen alfa hata ve örnek boyunda mevcut olan rasgele değişkenliğinin miktarına bağlıdır.

Örnek:
"İstatistik analizleri fibromiyalji ve tiroid otoimmünitesi arasında bir bağlantı olduğunu gösterdi:
OR = 3.8, % 95 GA = 1.54-10.13"
(Ribeiro RS, Proietti JA 2004)
Varsayalım ki p değeri = 0.03

BUNUN ANLAMI ŞUDUR:
Fibromiyalji ve tiroid otoimmünitesi arasında bir bağlantı yoksa, (şans eseri) 3.87 yada daha aşırı bir OR bulma ihtimali 0.03'tür

OR = 3.87, % 95 GA = 1.54-10.13

BUNUN ANLAMI ŞUDUR:

- Çalışma fibromiyalji ve tiroid otoimmünitesi arasında bir bağlantı olduğunu düşündürmüştür, yani, saptanabilir antitiroid peroksidaz antikorları ve/veya antitiroglobülin antikorları olan kişilerde fibromiyalji olasılığı diğer kişilerdekinden 3.87 kat daha fazladır.
- Yüzde 95 güven ile, gerçek odds ratio 1.54 ve 10.13 arasında (herhangi bir yerde) bulunur. Aralığın sıfır değerini (OR = 1.0 içermediği düşünülürse, gözlemlenen bağlantı alfa = 0.05 düzeyinde istatistik açısından anlamlı DEĞİLDİR.

KARIŞTIRICI

- Etkilerde karışıklık
- Konu dışı bir faktörün etkisi gerçek maruz kalma etkisiyle (sıfır olabilir) karıştırıldığından ilgili maruz kalmanın belirgin etkisi karışır

KARIŞTIRICI FAKTÖRLER İÇİN KRİTERLER

- Karıştırıcı faktörler hastalık için risk faktörü olmalıdır (hastalığın gerçek nedeni ya da gerçek nedeni için marker olabilir)
- Karıştırıcı faktör kaynak popülasyonda (vakaların kaynaklandığı PAR) araştırılan maruz kalmayla bağlantılı olmalıdır

KARIŞTIRICI FAKTÖRLER İÇİN KRİTERLER

Karıştırıcı fatör maruz kalma ya da hastalıktan ETKİLENMEMEKTİR. Konu dışı bir risk faktörü olmamalıdır - özellikle, maruz kalma ile hastalık arasındaki nedensellik zincirinde bir basamak OLMAMALIDIR (ara değişken değil)

arasında değişir". Fark istatistik açısından anlamlı olup, her ne kadar bizim bulduğumuz 3.8 oranı başka çalışmalarda farklı bulunabilse de, bu çalışma benzer şartlarda 100 kez tekrar edilse 95 keresinde elde edilecek OR değeri 1.54 ile 10.12 arasında olacaktır.

Burada, yeri gelmişken *odds* (tahmini risk) kavramını hatırlatmak isterim. Başka bir deyişle, risk olasılıktan farklıdır. Olasılık bir toplum içindeki payı ifade eder ama *odds*, olma durumunun olmama durumuna oranıdır. Yani bu örnekte, fibromiyalji olmanın fibromiyalji olmamaya oranıdır. Tüm grup içindeki fibromiyalji denirse olasılık ya da riskten söz edebiliriz. Türkçe'de *odds ratio*'yu tahmini rölatif risk olarak kullanıyoruz ama bunu riskten ayırt edebilmek çok önemlidir.

Kısaca karıştırıcı faktörleri de tanımlayalım. Karıştırıcı faktör üzerinde çalıştığımız etken ve sonuç ilişkisinden farklı üçüncü bir faktördür, ancak karıştırıcı faktör olmanın belirli diğer kuralları da vardır. Birincisi, karıştırıcı faktörün hastalıkla ilişkisi olmalıdır ama nasıl bir ilişki? Ya iki yönlü bir ilişki olacaktır ya da karıştırıcı faktör hastalığa yol açıyor olmalıdır. Karıştırıcı faktör hastalıktan bağımsız olarak etkenle de ilişkili olmalıdır. Benzer şekilde, bu ilişki ya iki yönlüdür ya da bu etkenle karşılaşmayı engelleyecek ve dışsal bir faktör olmalıdır.

Karıştırıcı faktör etkenle hastalık arasındaki ilişkiden bağımsız olarak bu durumu bozabilir, olmayan bir ilişkiyi varmış gibi gösterebilir, var olan bir ilişkiyi olduğundan az ya da çok gösterebilir. Son olarak, karıştırıcı faktör etkenin etki mekanizması üzerinde (dahilinde) rol oynayan bir ara faktör değildir.

Dilerseniz yukarıdaki örnek üzerinde çalışalım: İleri yaşta daha çok inme görülür, ileri yaşta saçlar ağarabilir; dolayısıyla kır saçlı olmakla inme arasında ilişki olabilir. Benzer şekilde yaş, çocuk sayısıyla meme kanseri arasında bir ilişki varmış gibi gösterebilir ya da kiralık evde oturanlarda koroner arter hastalığına ev sahibi olanlardakine göre daha sık rastlanıyormuş gibi görünebilir.

Bir faktörün çalıştığımız etken sonuç ilişkisini değerlendirmede karıştırıcı olabileceği düşünülüyorsa iki şey yapılır. *Adjusted* (uyarlanmış), yani o faktör için kontrol edilmiş değerler, bir de çalışmada karıştırıcı faktör yok sayılarak incelenen etken-sonuç ilişkisine ait kaba değer bulunur ve karşılaştırılır.

KARIŞTIRICI FAKTÖR NASIL DEĞERLENDİRİLİR?

- Karıştırıcı olarak sınanan faktörde tabakalandırma
- Etki boyutunun tabakaya spesifik tahminlerinin karşılaştırılması ve zıtlıklarının belirlenmesi
- Tabakaya spesifik etki boyutları kabul edilebilir derecede yakınsa uyarlanmış etki tahminleri hesaplanır
- Uyarlanmış değerler kaba değerlerle karşılaştırılır

KABA VE UYARLANMIŞ KARŞILAŞTIRMASI

	Hastalık +	Hastalık -	TOTAL
Maruz kalma +	a_i	b_i	E_{1i}
Maruz kalma -	c_i	d_i	E_{0i}
TOTAL	D_{1i}	D_{0i}	

- OR= a.d/b.c
- Uyarlanmış OR $\Psi_{MH} = \sum \frac{a_i d_i}{n_i} \Big/ \sum \frac{b_i c_i}{n_i}$
- PR= $\sum \frac{a_i d_i}{c/c+d}$
- Uyarlanmış RR $\Psi_{MH} = \sum \frac{(a_i E_{0i}/n_i)}{n_i} \Big/ \sum (c_i E_i/n_i)$

KARIŞTIRICI FAKTÖR NASIL DEĞERLENDİRİLİR?

FAKTÖRÜN ETKİ DEĞİŞTİRİCİ OLMADIĞINDAN EMİN OLUN!

- Etki boyutunun kaba yada uyarlanmış tahminlerini karşılaştırın:
  - tahminlerdeki değişiklik yüzdesiyle (genellikle % 10-15)
  - hipotez testiyle

Etki Değişimi (Etkileşim)

- Antagonizm
- Sinerji
- Toplamsal etkileşim
- Çarpımsal etkileşim

→ ETKİNİN (BOYUTUNUN) TABAKAYA SPESİFİK TAHMİNLERİ ANLAMLI DERECEDE FARKLIDIR

↓

AÇIKLAYIN!

A Genindeki mutasyon ve KKH arasında bir bağlantı var mı?

	KKH +	KKH –	Total
Mutasyon +	400	1600	2000
Mutasyon –	100	1900	2000
Total	500	3500	4000

$$PR = \frac{400/2000}{100/2000} = 4$$

Irk A Genindeki mutasyon ile KKH arasındaki bağlantıyı nasıl etkiliyor?

	Kafkasyalılar			Afrika kökenli Amerikalılar	
A Geninde mutasyon	KKH (+)	KKH (+)	Risk 0.060	KKH (+)	Risk (-)
var	24	376	0.060	1224	0.235
yok	96	1504	0.060	396	0,010
RR =			1.0		23.5

İstatistikçiler çoğu durumda bu ikisi arasındaki farka bakıp anlamlı olup olmadığını söyler, yani formel bir hipotez testi yapar ve testin p değerine göre karar verirler. Epidemiyologlar kaba ve uyarlanmış değerleri karşılaştırmaya ek olarak, biraz daha farklı bir şey yapar: bu formel hipotez testi olabilir ya da nokta tahminde %15 ten fark varsa anlamlı olduğu düşünülebilir. Faktörün karıştırıcı faktör olması için sahip olması gerekli 3 özelliğin olup olmadığı incelenir.

Vurgulamak istediğim nokta karıştırıcı faktörün gerçekte olmayan bir şey olduğudur. Veri seti özelinde vardır ve çalışılan etken-sonuç ilişkisini yorumlamada karışıklığa neden olur. O zaman çalışmada bunu kontrol etmek gerekir.

Etkileşim (*interaction*) veya etki değişimi (*effect modification*) ise gerçekte olan bir durumdur. Öyle ki, bazı faktörlerin alt gruplarında etken-sonuç ilişkisi farklıdır. Bu durumda, bu faktörü kontrol etmek doğru değildir, aksine bulunan farklı sonuçlar ayrı ayrı bildirilmelidir. Bir faktör kadınlarda koruyucuysa ve erkeklerde de riskse bundan korunulması söylendiğinde kadınlara kötülük yapılır. O zaman kadınlar ve erkekler için farklı yorum yapmak gerekir. Ters etkileşim ya da aynı yönde etkileşim olabilir. Aynı yönde etkileşirken (1) toplamsal bir etkileşim olabilir, yani bir faktör iki kat, öbürü üç kat riske neden oluyorsa toplam beş kat olur, ya da (2) çarpımsal bir etkileşim olabilir, altı ve daha üstü kat bir ilişki saptanır.

Örneğin, koroner arter hastalığıyla ilgili A geni adlı bir gen olduğunu varsayalım. Bu gende bir nokta mutasyonu bulundu ve bu durumun daha çok koroner kalp hastalığına neden olduğu iddia ediliyor. Tabloya baktığınızda, 2000 mutasyonlu kişinin 400'ünde koroner kalp hastalığı olduğunu, 2000 mutasyonu olmayan kişinin 100'ünde koroner kalp hastalığı olduğunu görüyorsunuz. Bu mutasyonun olduğu durumlarda 4 kat risk vardır; başka bir deyişle, koroner arter hastalığı 4 kat daha fazla gözlenir. Ancak, bu ilişki her ırkta geçerli olmayabilir. Örneğin, bu çalışma Amerika'da gerçekleştirildiyse, beyaz ırkta ve Afrika kökenli Amerikalı'larda araştırma yapılmış olur. Beyaz ırkta riskin 1 olduğunu, beyaz ırkta hiç fark yaratmadığını görüyorsunuz ama siyah ırkta risk 23.5'tir. Sonuçta bu genin 4 kat riskli olduğu, bundan korunmak için bir şey yapmak gerektiği söylenebilir.

Ancak, bu durumda, beyaz ırk için yanlış bir karar vermiş olunur (zaten risk yoktur). Öte yandan, siyah ırkın riski de hatalı olarak az söylenmiş olur.

Bu örnekte bir etkileşimden söz edilir. Yani, ırkın alt gruplarında gen-koroner arter hastalığı ilişkisi için farklı sonuçlar alınır ve bunun şöyle bildirilmesi gerekir: Beyaz ırkta risk yok, siyah ırkta ise 23.5 kat fazla risk vardır. Bu aşamaya gelindiyse artık karıştırıcı olup olmadığına bakılmamalı, etken-hastalık ilişkisi ayrı ayrı bildirilmelidir. Etkileşim bulunmazsa, o durumda ırkın karıştırıcı faktör olarak rol oynama olasılığını incelemeye devam etmek gerekir.

O halde, bu araştırmada etken-sonuç ilişkisini araştırırken, ırkın rolü için ilk bakılacak özellik, ırkın etken ile etkileşimi olup olmadığı, beyaz ırk ve siyah ırkta farklı olup olmadığıdır. Beyaz ve siyah ırkta etken-sonuç ilişkisi farklı değilse o zaman ırkın karıştırıcı faktör olma olasılığına bakılmalıdır. Karıştırıcı faktör de değilse, olduğu gibi, 4 değeri verilmelidir. Farklıysa, ırk için uyarlanmış (*adjusted*) değer verilir. Gruplara özel değerler farklı değil, beyaz ırkta da siyah ırkta da aynıysa, örneğin, ikisinde de 3 ya da 2,5 olmasına rağmen sonuç 4 çıkmışsa, etkileşim olmadığı kabul edilir, karıştırıcı faktör açısından kontrole geçilir.

Son bir örnek daha değerlendirelim. Çalışmanın hipotezi "genç yaşlarda kahveye gidenlerde 50 yaşından sonra koroner kalp hastalığı gelişiyor" olsun. Verilere baktığımızda, bir grupta 110 kişinin kahveye gittiğini ve 81'inde koroner kalp hastalığı geliştiğini, ikinci grupta 210 kişinin kahveye gittiğini ve 28'inde kalp hastalığı olduğunu görüyoruz. Yani kalp hastalığı riski grupların birinde %73, diğerinde %13'tür. Dolayısıyla, kahveye gittiyseniz koroner kalp hastası olma riskiniz 5.5 kat fazladır. Burada ilk akla gelen kahvede sigara dumanı olmasıdır çünkü pasif sigara içiciliği koroner kalp hastalığında çok önemlidir.

Peki, bunun cinsiyet ile ilişkisi olabilir mi? On kadın kahveye gitmiş ve sadece biri koroner kalp hastası olmuş, 200 kadın gitmemiş ama onların arasında da 20 kişi koroner kalp hastası olmuştur. Dolayısıyla, koroner arter hastalığı olma riski kahveye gidilse de gidilmese de benzerdir. Yani rölatif risk 1'dir, başka bir deyişle yoktur.

Gençken boş zamanlarda kahveye gitmek gelecekte KKH açısından risk faktörü müdür? (20 yıl izlemeli retrospektif kohort çalışması)

	KKH +	KKH-	Total
Giden	81	29	110
Gitmeyen	28	182	210
Total 1	109	211	320

Kahveye gidenlerde KKH riski = % 73.6
Kahveye gitmeyenlerde KKH riski = %13.3
RR = 5.53

Boş zamanlarda kahveye gitme?

Kahveye gitme	Kadınlarda			Erkeklerde		
	KKH (+)	KKH (-)	Risk	KKH (+)	KKH (-)	Risk
Evet	1	1	0,1000	80	20	0,8000
Hayır	20	180	0,1000	8	2	0,8000
RR =			1.0			23.5

- Gençlikte boş zamanlarda kahveye gitmek (E) erkeklerde de (R= 1.0) kadınlarda da (R= 1.0) gelecekte KKH ile bağlantılı DEĞİLDİR
- Cinsiyet (C) E ile bağlantılıdır, yani 100/110 (%90,9) erkek ve 10/210 (%4.8) kadın gençliğinde kahveye gitmiştir

  C ⟶ E
- Cinsiyet (C) KKH (D) ile de bağlantılıdır, yani 88/100 (% 88.0) erkek ve 21/210 (10.0) kadın gençliğinde kahveye gitmiştir

  C ⟶ D

C ⟶ E KABA RR = 5.53 vs.
C ⟶ D Uyarlanmış RR= 1.0

KARIŞTIRICI FAKTÖRÜN KONTROLÜ
BİR ÇALIŞMA PLANLANIRKEN:

- Kısıtlama (dışlama; spesifikleştirme) iç geçerliliği artırır
- Eşleştirme
- Randomizasyon (klinik çalışmalarda)

KARIŞTIRICI FAKTÖRÜN KONTROLÜ
ANALİZDE:

- Standardizasyon
- Tabakalı analiz - Mantel - Haenszel OR, RR kullanılır
- Çok değişkenli teknikler (çok değişkenli lojistik regresyon)

Riskin 1 olması demek, etkenle karşılaşan (burada etken kahveye gitmektir) ve karşılaşmayanların aynı riske sahip olması demektir.

Peki erkekler? Erkekler gitse de gitmese de % 80 koroner kalp hastası oluyorlar. Onlarda da risk 1'dir. O zaman ne yapılmalıdır? Bunlara etki değişimi olsaydı, örneğin, erkeklerde 5, kadınlarda 1 olsaydı, erkeklerin kahveye gitmemesi gerektiği, kadınların ise riski olmadığı için kahveye gidebileceği söylenebilirdi, ancak etkileşim yoktur. Öte yandan, kaba (*crude*) değerin 5.5, uyarlanmış (*adjusted*) değerin ise 1 olduğunu görüyoruz. Bir ile 5.5 farklıdır ancak farklı demek, cinsiyet karıştırıcıdır demek için yeterli değildir.

İnceleyeceğimiz bir konu da karıştırıcı faktörün (cinsiyet) hastalık ile ilişkili olup olmadığıdır. Erkeklerin % 90'ı, kadınların ise % 5'i koroner kalp hastasıdır. Arada çok büyük fark görüldüğüne göre, cinsiyet koroner kalp hastalığıyla ilişkilidir. Özetle, cinsiyetle koroner kalp hastalığı arasında erkekler aleyhine bir ilişki mevcuttur; başka bir deyişle erkekler daha çok hasta olmaktadır.

İkinci sorumuz şudur: karıştırıcı faktör, hastalıktan bağımsız olarak etken ile (etken kahveye gitmek) ilişkili midir? Erkeklerin % 90'ı, kadınların ise % 5'i kahveye gittiği için ilişkilidir. Koroner kalp hastalığı erkeklerin % 88'inde, kadınların ise % 10'unda mevcuttur. Sonuç olarak, kahveye erkekler gittiği için ve erkeklerde daha çok koroner kalp hastalığı olduğu için, aslında kahve ile koroner arter hastalığı arasında hiçbir ilişki olmadığı halde, sanki kahveye gitmekle koroner kalp hastalığı arasında bir ilişki varmış gibi görünür.

Son olarak, üçüncü faktör nedir? Karıştırıcı faktör, etkenin sonuç ile olan ilişki yolu üzerinde (arada) bir değişken değildir. Kahvehaneye gitmekle koroner arter hastalığı arasındaki ilişkide cinsiyet bir ara basamak olabilir mi? Hayır. O halde cinsiyet bu çalışmada karıştırıcı faktör olarak rol oynamaktadır ve kontrolü gerekir.

Karıştırıcı faktörler çalışmayı planlama aşamasında kontrol edilebilir. Planlamada bu değişkenlerin olası karıştırıcı etkilerinden haberdar olunmayabilir veya kontrol sağlanamaz ise, analiz aşamasında da kontrol edebilir. Bu özellik, araştırmacının lehine sonuçlanabilir.

Kontrol için planlama aşamasında öncelikle kısıtlama yapılabilir, bir çalışmanın sadece erkeklerde yapılacağı, erkeklerin bildirileceği söylenebilir. İkincisi, eşleştirme yapılabilir. Erkeklerle kadınlar arasında bir fark varsa eşleştirilebilir; her erkek için bir erkek, her kadın için bir kadın kontrol alınabilir ve böylece fark yaratanın erkek mi, kadın mı olduğu önemsenmeyebilir.

Bir önemli konu da, klinik çalışmalardaki randomizasyondur. Klinik çalışmalarda müdahale grupları ve plasebo da etkendir. Mevcut ilaçla karşılaştırırken randomizasyon yapıldığı zaman, bilinen ve bilmeyen tüm değişkenler (karıştırıcı faktörler) için kontrol sağlanır.

Planlamada kontrol sağlanamadığı durumlarda karıştırıcı faktör kontrolü istatistik analizi sırasında da yapılabilir. Örneğin, ülkelerdeki sigara içme-kanser ilişkisi araştırılırken yaş için standardizasyon yapılabilir. İkinci bir seçenek de, tabakalı analiz yapıp, tabakalara özel nokta değerlerden yola çıkarak (örneğin, Mantel, Haenszel *odds ratio* değeri) veya ağırlıklandırılmış ortalama değerleri kullanılabilir. Son olarak, lojistik regresyon benzeri bir çok değişkenli analiz yöntemi kullanılarak, birden çok karıştırıcı faktörün etkisi aynı anda kontrol edilebilir.

KATKILAR

John P. A. Ioannidis: İstatistik, klinik ve biyolojik anlamlılık arasında fark vardır ve bir çalışmada her üçünün de bir arada bulunması gerekmez. Klinisyenlerin dediği gibi bir çalışmada önce klinik ve ardından onu destekleyecek biyolojik, en sonra da istatistik kuvvetin önemi vardır. Buraya kadar tüm söylenenlere tümüyle katılıyorum. İstatistik anlamlılık konusuna gelince, burada çoğu kez şu yanlışa düşüyoruz. Dünyada tek bir araştırıcı, tek bir çalışma, sınanan tek bir hipotez varsa, yapılan çalışmadan başka hiçbir çalışma planlanmıyor ise, tek bir istatistik analiz yapılacak ise ve de tüm bunlar hiçbir yanlılık, karıştırıcılık içermiyorsa, o zaman elde edilen p değeri çok güvenlidir. Ancak çalışma sayısı, varsayım ve istatistik analiz birden fazla ise o zaman işler değişiyor.

P değerleriyle ilgili şimdiye kadar çok sürprizle karşılaştım. Genellikle yaptığı şey şudur: bir çalışmayı p değerinden çok geçerliliği ile değerlendirmeyi her zaman yeğliyorum.

P değeri ise bir geçerlilik göstergesi değildir. Bazen aynı veriler üzerine o kadar çok analiz yapılabilir ki, araştırıcı bile kaç kez analiz yaptığını unutur veya bir araştırma gayet güvenilir olur, ancak başkaları bu araştırma verilerine girip anlamlılık testleri yapmaya başlarlar. Böyle durumlarda da anlamlı gibi görünen p'ler ortaya çıkabilir. O nedenle bir araştırmayı planlarken araştırıcının önce kendine sorması gerekir: Bu araştırmayı esasında neden yapıyorum? Gerçekte kaç tane varsayım sınıyorum? Gerçekten gerçeği mi arıyorum? Bu araştırmayla bir takım ilişkiler mi arıyorum yoksa esas amacım tek bir varsayımı mı sınamak?

Bu arada belirtmeliyim, randomize kontrollü çalışmalarda bile karıştırıcı faktörler daima vardır. Randomizasyonla bile gruplamaların yanlılık olmadan seçildiğine %100 emin olunamaz. Karıştırıcılığı ortadan kaldırmak için kullandığımız istatistik yöntemler o kadar güçlü değildir. Bunun nedeni şudur: karıştırıcı, sonuç ve karıştırıcının etkilediği grup arasında yapısal ilişkiler göz önüne alınmaz. İşler göründüğünden daha çapraşıktır. Ortada bir karıştırıcı görülmediği durumlarda bile bu çalışmada karıştırıcı olmadığını söylemek çok zordur. Özetle karıştırıcıdan kurtulmak hiç kolay değildir.

Banu Çakır: Karıştırıcılıkla ilgili bir şey eklemek isterim. Kontrollü klinik çalışmalarda randomizasyon, bilemediğimiz karıştırıcıların, çalışma grupları arasında olabildiğince eşit olarak dağılmasını sağlar. Sonuçları analiz ederken ortaya çıkan karıştırıcı faktörleri de çok katmanlı analizlerle ortadan kaldırmaya çalışırız. Ancak herhangi bir çalışmada tüm potansiyel karıştırıcıları ortadan kaldırdığımıza tümüyle emin olamayız. Klinisyenler bazı tıp dışı faktörlerin de karıştırıcı olabileceğini unutmamalıdır. Örneğin bir hastaya ilaç yazıldığında hastanın o ilacı alıp almayacağına emin olunamaz. Hastalar kendilerine yazılan ilaçları uzun süre almak zorundadırlar. Hastanın sosyoekonomik durumu, hastanın aldığı bakımın kalitesi, hastanın sağlık hizmetlerine ulaşabilme olanakları veya sadece hastanın uyumlu bir kişiliği olup olmadığı karıştırıcı faktörler olabilir. Dr. Ioannidis'in dediği gibi bazen bunlar kontrol edilemez, bazen ortada bir karıştırıcı olabileceği akla bile gelmez. Bu durumlarda, bulunan bir ilişki geçersiz olabilir. Bunun yanında ortada bir karıştırıcı olması gerçekte olan

bir ilişkiyi maskeleyebilir veya bulunan bir ilişkinin derecesini etkiler. İşte bu nedenlerden de gerçek bilimsellikte sonlanmış sanılan çalışmaları tekrar tekrar gözden geçirmemiz, gerekirse daha iyi yöntemlerle tekrar yapmamız ve yetişen gençlere de bu görevleri aşılamamız gerekir.

KAYNAKLAR:

1. Bellhouse DR. Invited Commentary: p values, hypothesis tests, and likelihood. Am J Epidemol 1993;137:4979.
2. Çakır B. Tıbbi Makale Okuma, Hazırlama ve Değerlendirmede Epidemiyolojik Yöntemlerin Kullanımı. Sağlık ve Toplum Dergisi 2003;4:39.
3. Çakır, B. Do not get confused by the confounders: identification and control of confounders in medical research. Journal of Public Health 2004;2(1):3443.
4. Çakır B. Sağlık Araştırmalarında Bias (Yanlılık, Taraf Tutma): Tipleri, sınıflandırılması, nedenleri, önleme yöntemleri I. Türkiye Klinikleri Tıp Bilimleri Dergisi. 2005;25(1):10010.
5. Çakır B. Sağlık Araştırmalarında Bias (Yanlılık, Taraf Tutma): Tipleri, sınıflandırılması, nedenleri, önleme yöntemleri. **II.** Türkiye Klinik leri Tıp Bilimleri Dergisi. 2005;25(2):2808.
6. Goodman SN. P values, hypothesis tests, and likelihood: Implications for epidemiology of a neglected historical debate. Am J Epidemiol 1993;137:48596 and 5001, and Comment in 1994;139(1):1168.
7. Kalsbeek W, Heiss G. Building bridges between populations and samples in epidemiological studies. Annu Rev Public Health. 2000;21:14769.
8. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. Van Nostrand Reinhold, New York, NY, 1982.
9. Last JM (ed.). A dictionary of epidemiology. International Epidemiology Association. 4th ed. Oxford University Press, New York, NY , 2001.
10. Rothman K, Greenland S. Modern epidemiology. Second ed. LippincottRaven Publishers, Philadelphia, PA, 1998.
11. Hernan MA, HernandezDiaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol 2002;155:17684.
11. Weinberg CR. Toward a clearer definition of confounding. Am J Epidemiol 1993;137:18.

BİR DEMET YANLILIK

Hasan Yazıcı

Profesör Doktor, İstanbul Üniversitesi, Cerrahpaşa Tıp Fakültesi, Romatoloji Bilim Dalı

Ben *bias*'ın Türkçe karşılığı olarak iki ayrı sözcük önereceğim. Biraz sonra değineceğim gibi *bias* kasıtlı veya kasıtsız olabilir. Önerim kasıtlı *bias*'a yanlılık, kasıtsız *bias*'a da saptıran demek. Ancak bu dediğim genel kabul görür mü bilmiyorum. O nedenle konuşmamda herhangi birini kullandığımda şimdilik diğeriyle eşanlamlı düşününüz.

Her şeyden önce konuyla ilgili bazı kavramların üzerinden bir daha geçmek istiyorum. Tablo 1'de bu kavramların tanımlarını, İngilizce karşılıklarıyla beraber sıraladım.

Ana kavramları böyle tanımladıktan sonra artık yanlılık/saptırma'nın (*bias*) da tanımını yapabiliriz. Yanlılık/saptırma geçerlilikten uzak düşmektir. Yukarıda değindiğim gibi, işin içinde kasıt varsa *bias*'a en uygunu yanlılık demektir. Olay kasıtsız ise (örneğin ölçü aletimizde haberimiz olmayan bir yanlış gösterme varsa) saptırma denir ve değişkenlik çoğu kez saptırmanın önemli bir öğesidir.

Tablo 1.

- Doğruluk, geçerlilik, (*validity*) : gerçeğe yakınlık
- Tekrarlanabilirlik (*reproducibility, reliability*) : daha evvelki ölçülere yakınlık
- Değişkenlik (*variability*): tekrarlanabilirlik eksikliği
- Hata (*error*): doğrudan uzaklaşmak
- *Truth* (doğru) = Gerçek

Tablo 1'de sıraladığım kavramlar arasında unutulmaması gereken önemli bir nokta da tekrarlanabilirliğin geçerliliğin bir öğesi olmasıdır. Ancak tekrarlanabilirlik geçerlilik için yeterli bir öğe değildir. Tıpta en sık yapılan ölçme yanlışlarından biri salt tekrarlanabilire geçerli demektir. Örneğin laboratuvarda kolesterolünüzü ölçtürüyorsunuz. Diyelim ki gerçek kan kolesterol düzeyiniz 200 mg/dL. Söz konusu laboratuvarda kolesterolünüz ilk ölçüldüğünde 230mg/dL çıksın. Ölçümü yapana dönüp "bu doğru mu (= geçerli mi)?" diye sorarsanız laborantın (ve korkarım laboratuvar sorumlusu doktorun da) yapacağı ilk şey sizden aldığı örneğe aynı makinada bir kez daha bakmak olacaktır. Eğer bu değer ilk ölçüme çok yakın, diyelim 231mg/dL çıkarsa sorunuza alacağınız yanıt büyük olasılıkla, "Hocam, bu ölçtüğümüz doğru," olacaktır. İşte burada yapı-

lan klasik yanlış, tekrarlanabilire doğru (= geçerli) demektir. Doğal olarak, makbul olan hem doğru hem de tekrarlanabilir olan ölçüdür.

Yanlılık veya saptırma dediğimiz zaman ise sistemik, kasıtlı veya kasıtsız bir geçerlilik eksikliğini anlıyoruz. Hepimizin içli dışlı olduğu olgu şudur: kontrol çalışmaları saptırmaların en sık olduğu yerdir, ancak saptırmalar kesinlikle bununla sınırlı değildir. Laboratuvar çalışmasında, ilaç çalışmasında da mevcuttur. Örnek olarak randomize kontrollü ilaç çalışmalarını alalım. Randomize kontrollü çalışmada daha evvel çalışılacak ilacı almış ve yanıt vermemiş insanlardan seçip o insanlara yeni ilaç yanında bu eski ilacı da vermeye devam edersek, çalışmanın sonunda yeni ilacın eski ilaca üstün çıkacağı kesindir.

Yanlılığı/saptırmayı genellikle üçe ayırıyoruz:
Yanlılık/Saptırma (*bias*) çeşitleri

- Bilgi toplama (*information*)
- Seçme (*selection*)
- Karıştırıcılık (*confounding*)

I. Bilgi toplama yanlılığı:

a. Gözlemcinin niteliklerine
b. Deneğin niteliklerine veya
c. Gözlemleri yaptığınız araç ve/veya yöntemin niteliklerine bağlı olabilir.
Yanlılık bilgiyi toplamakla ilgili (bu sizdeki bir takım sorunlardan olabilir), denekteki sorun larla ilgili veya ölçtüğünüz aletten dolayı ola bilir. Yanlılıktan söz ederken geçerlilik kavramına kısa bir dönüş yapayım.
Geçerliliği ikiye bölmek gerek:

1. İç geçerlilik veya tutarlılık (*internal validity*)
2. Dış geçerlilik – veya genellenebilirlik (*generalizability*)

İç geçerlilikten kastımız, yapılan ölçümün gerçekten ölçmek istediğimizi ölçüp ölçmediğidir. Örneğin eritrosit sedimantasyon hızı romatoid artritin şiddetini ölçmekte iç geçerliliği oldukça kısıtlı bir ölçü yöntemidir. Sedimantasyon hızı yüksek olup da hafif romatoid artriti olan birçok hasta vardır. Bunun yanında klinik olarak hafif romatoid artritli bir grup hasta arasında denediğimiz bir ilacın, hastalığı ağır seyreden bir gruba da iyi geleceğini savunmak ise dış geçerliliği pek olmayan bir savdır. Bir de biraz önce Dr. Banu Çakır'ın da değindiği gibi, karıştırıcı ya da *confounding* dediğimiz sorunlar mevcuttur. Ben bunlara ayrı ayrı ve özellikle romatolojiden örnekler vermeye çalışacağım.

Önce bilgi toplamayla ilgili yanlılıklardan bahsedelim. Bunun en basit örneği şudur: diyelim ki kan şekeri ölçmek istiyoruz ve elimizdeki ölçü aleti kan şekerini devamlı olduğundan 30mg/dL düşük gösteriyor. İşte o zaman ortada doğrudan elimizdeki ölçü aletinin işlevini düzgün yapmamasından doğan bir yanlılık olur. Başka bir örnek daha vermek isterim. Diyelim ki bir grup insana geçen yıl içinde belirli bir nedenle herhangi bir hekime başvurup vurmadıklarını soralım. Deneklerimizin bir bölümü hekime gitmiş olsalar dahi bunu hatırlamayacaklardır (hatırlama yanlılığı veya *recall bias*).

Bilgi edinme yanlılığın bir üçüncü türü çift organlarda yapılan gözlemlerle ilgilidir. Örneğin, romatoid artritli bir hastanın el filmlerine bakarken sağ elde erozyon görülmüşse, sol elde erozyon ararken daha yanlı davranılır. Radyografide görülen, doğası esasında tartışmalı olması gereken bir lezyon, daha rahatlıkla erozyon diye yorumlanabilir. Çift organlar hatasına çok güzel bir örnek yıllar evvel ABD'de yaşayan bir Türk biyofizikçisi (El Cantekin) tarafından verilmişti. Burada konu çocuklarda orta kulak infeksiyonlarını izleyen hekimin çocuğun bir kulağına baktıktan sonra ikinci kulağına bakarken, birinci kulakta gördüklerinin etkisi altında kalmasıydı.

II. Seçme yanlılığı

Araştırmalarda en sık rastlanan yanlılık çeşidi seçme üst başlığı altında topladığımız yanlılıklardır.

Seçme (*selection*) türü yanlılık çeşitleri:

- Yetersiz eşleştirme (*Undermatching*)
- Aşırı eşleştirme (*Overmatching*)
- Berkson yanlılığı (*Berkson's bias*)
  komorbidite (*comorbidity*), klasik (*classic*)
- Hastalık başlangıç zamanı (*Lead time*)
- Hastalık süresi (*Disease duration*)
- Ekoloji (*Ecology*)
- Ortalamaya regresyon (*Regression to the mean*)

Yetersiz eşleştirme: Bir çalışmada araştırma grubunda aradığınız bir özelliğin, bu grubun zaten doğası gereği, kontrol grubunda daha az sıklıkta görülmesidir. Biraz evvel randomize kontrollü ilaç çalışmalarından verdiğimiz örneğe dönelim. Herhangi bir hastalıkta yeni bir ilaç (X)

denemek istiyorsunuz. Elinizde o hastalıkta geleneksel olarak kullanılan (Y) ilacını alan ama buna arzu edilen düzeyde yanıt vermeyenlerden oluşan bir hasta grubu var. Şimdi, bu hasta grubunu Y ilacını almaya devam etmek koşuluyla plasebo ekleyeceğiniz bir grup, A kolu, ve yine Y ilacını almaya devam ederken X ilacını ekleyeceğiniz B kolu diye iki kola ayırıp bir süre izliyorsunuz. Bu koşullarda, büyük olasılıkla hastalık iyileşmesi B grubunda A grubuna göre daha fazla görülecektir. Diğer bir deyişle bu çalışmada X ilacının Y ilacına üstünlüğü var dersek, yetersiz eşleştirmeye dayanan, geçerliliği az bir sonuca varmış oluruz. X ilacının Y ilacına gerçek üstünlüğünün sınanmasının bilimsel olarak doğru yolunun, çalışmanın daha önce ne Y ne de X ilacını almamış bir hasta grubunu X ve Y ilaçlarını alacak iki kola ayırıp gözlemek olduğu açıktır.

Aşırı eşleştirme: Burada ise durum tersidir. Örneğin ankilozan spondilitle (AS) ürogenital infeksiyon arasında yakın bir ilişki olduğu yolunda bir varsayımımız olduğunu düşünelim ve bu amaçla bir olgu kontrol çalışması planlamaya çalışalım. Yüz tane AS'li erkek hastamız var. Bu gruba kontrol grubu olarak, bir üroloji kliniğine dizüriyle başvuran 100 ardışık erkek hastayı alıp, her iki grup arasında prostatit sıklığına bakarsak, AS'liler arasında aradığımız artmış prostatit sıklığını belki hiç gösteremeyebiliriz. Bunun nedeni de üroloji kliniğinden aldığımız kontrol grubunda prostatit sıklığının tanımı üzere yüksek bulunacağıdır. AS'de prostatit olup olmadığını sınamak için kontrol grubu seçiminde kadın doğum kliniğine gitmek ne kadar yanlışsa üroloji kliniğine gitmek de o kadar yanlış olur!

Berkson yanlılığı: Berkson Mayo Clinic'ten bir istatistikçidir. Geçen yüzyılın ortasında, hastanelerde yapılan olgu kontrol çalışmalarından çıkan sonuçlarda, grupların bir hastaneye veya hekime başvuruş farklarından doğan sorunlar olduğunu anlamıştır. Berkson yanlılığı adı altında iki kavram mevcuttur: Birincisi genel komorbidite hatasıdır; bir insanın A hastalığı varsa bir hekime veya hastaneye başvurma olasılığı aynı hastada B hastalığı da olursa artar. Berkson Paradoksu ise daha değişik bir kavramdır. Örneklerle iki kavram arasındaki farkı açıklamaya çalışacağım:

FMF ile Behçet hastalığı arasında bir ilişki olduğu iddiası var. Bunu kanıtlamak isteyen bir araştırıcı FMF'li hastalar arasındaki aft sıklığını, örneğin, romatoid artritliler arasındaki aft sıklığıyla karşılaştırıyor. Eğer FMF'liler arasında da aft artmış bulunursa bu, FMF'in Behçet'e yakın bir hastalık olduğuna kanıt oluşturacaktır. FMF'liler arasında topluma göre aft sayısının arttığını gösterirlerse varsayımları bir derece doğrulanacaktır. Peki ne yapıyorlar? Hastaların eşlerinin ağzında afta bakıyorlar. Bu örnekte eğer FMF'li hastaların ağızlarında eşlerinden daha fazla aft gözlenirse buna pek şaşırmamak gerek. Çünkü yukarıda bahsettiğim gibi bir insanda iki patoloji birden olursa (komorbidite) doktora veya hastaneye başvurma olasılığı artmaktadır.

Zaten sağlıklı bir insanda (elimizdeki örnekte eşler) FMF'lilerden daha az sıklıkta aft beklenir. O nedenle burada yapılması gereken, sağlıklılar yanında, başka hastalık gruplarında da aft sıklığını araştırmaktır. Ancak burada da Berkson yanlılığının klasik denilen şekline dikkat etmemiz gerektir.

Klasik Berkson yanlılığı: Görüldüğü üzere Berkson komorbidite yanlılığında sorun, kıyaslamanın hastalar (FMF'liler) ve bir hastaneye (veya hekime) başvurmayanlar (eşler) arasında yapılmasından doğuyordu. Klasik Berkson hatasında ise, hem hasta grubu ve hem de kontrol grubunun (grupları) hastaneye başvurması ve yanlılığın ortaya çıkış nedeninin söz konusu başvuru sıklıklarının farklı olmasıdır.

Örneğin, diyelim ki bir "A" ve bir "B" hastalığımız var. Bunlar arasındaki ilişkiye bakıyoruz. İki şey yapıyoruz: A'lar arasında B ve B'ler arasında A arıyoruz. Şu şekilde düşünelim: poliarteritis nodosa (PAN) hastaları arasında vitiligo görülme sıklığının, sedef hastaları arasında vitiligo görülme sıklığından daha fazla olduğunu düşünüyoruz. İşte bu amaçla bir hastanenin:

a. romatoloji polikliniğine gidip oraya gelen lupus hastaları arasında vitiligo tarar, ve
b. yine aynı hastanenin dermatoloji polikliniğine gidip oraya gelen sedef hastaları arasında vitiligo tararsak,

PAN'lı hastalar arasında vitiligo gerçekten de artmış olsa bile bunu kanıtlayamayız. Nedeni ise şöyle: PAN ağır bir hastalıktır. PAN'ı olan hastalar esas hastalıklarına ek vitiligo olsun ya da olmasın hastaneye başvuracaklardır. Halbuki toplumdaki sedef hastalarının PAN'a göre az bir kıs-

mı hastaneye başvurur. Ancak bir sedef hastasında vitiligo da olursa hastaneye başvurma sıklığı artar. İşte bu nedenle de hastaneye gelen sedef hastaları arasında vitiligo sıklığı gerçekte olduğundan artmış görülür.

Klasik Berkson yanlılığın işleyiş kuralını şöyle özetleyebiliriz:

I.A ve B hastalığı olanlar arasında C bulgusu nun (belirti veya hastalık da olabilir) sıklığını araştırıyorsak,

II.A hastalığı olanların salt A hastalığı nedeniyle hastaneye başvurma sıklıkları B hastalığı olanlara kıyasla daha fazla ise,

III.C bulgusunun kendisi hastaneye başvurmak için önemli bir neden değilse,

B hastaları arasındaki C sıklığı, A hastaları ara sında görülecekten daha fazla olacaktır.

Berkson yanılması bize "olgu kontrol çalışmaları yaparken seçtiğin kontrol gruplarının hastaneye başvurma sıklıklarını da göz önünde tut" diyor.

Hastalık başlangıç zamanı *(lead time)* yanlılığı: Bu yanlılığın temelinde bir hastalığa erken tanı koymanın her zaman yaşamı uzatmayacağı fikri yatar. Bu durumun klasik örneği de prostat kanseridir. Prostat kanserinde yaşam süresinin ortalama 20 yıl olduğunu farz edelim. PSA testi de, tuşe bulgularının ortaya çıkmasından ortalama 5 yıl önce tanı koydurabilsin. Kanseri erken evrede yakalamak gerçek bir şifa sağlayamıyorsa (ki prostat kanserinde böyle bir şifanın olduğu hala kuşkuludur) yüksek PSA ile 5 yıl erken tanı konan hasta ameliyat edildiğinde, ameliyat sonrası yaşam süresi erken tanı konan hastada, tuşe yöntemiyle tanı konana kıyasla 5 yıl daha uzun olacak ancak her iki hastanın da mutlak yaşam süresinde birşey değişmeyecektir.

Hastalık süresi *(Disease duration):* Diyelim ki bir hastalık (örneğin kuş gribi) insanları 1 ayda öldürüyor. Şimdi biz toplumda 3 ayda bir kuş gribi taraması yaparsak, bu sayı bize o toplumdaki gerçek kuş gribi sıklığını yansıtmayacaktır. Bunun nedeni de açıktır. Üç ayda bir o toplumda kuş gribi taranırsa, kuş gribi sıklığı pek fazla bulunamaz çünkü bir bölüm hasta, hastalığı tarama sıklığından daha kısa sürede hastalanıp, hastalık süresini tamamlayıp ölmektedir.

Ekoloji: Bu yanlılığa epidemiyolojik çalışmalarda sık rastlanır. Bir toplumda görece veya mutlak sık görülen bir özelliği o toplumu oluşturan bireylere de özgü sanmak yanlılığıdır. Örneğin ülkemizde tüberküloz ve Behçet hastalığı sıktır. Bundan tüberküloz ve Behçet hastalığı arasında bir ilişki olduğu sonucu çıkmaz. Nitekim tüberküloz geçirmiş ve geçirmemiş iki grup arasında Behçet sıklığı araştırıldığında iyi bir olasılıkla sıklıkların birbirine yakın olduğu görülecektir.

Ortalamaya regresyon: Eğer bir toplumdan en uç örnekleri, örneğin en uzun boyluları alırsak, aynı toplumdan ikinci alacağımız örnek o toplumun ortalamasına daha yakın olur. İlk kez 19. yüzyılda ünlü genetikçi Galton tarafından ortaya atılmış bir kavramdır. Bu ünlü örneğe göre çok uzun boylu babaların çocukları yine toplumun ortalamasına göre uzun boyludur, ancak babalarına göre ortalamaya daha yakındırlar.

McDonald çalışmasında, İskoçya'da bir hastanede, o hastaneye çeşitli tanılarla yatanların çeşitli laboratuvar değerlerine bakarak, üç standart varyasyonun üzerinde bulunan değerler alınmıştır. Aynı insanların aynı değerlerine bir taburcu oldukları sırada, bir de ilk yattıkları zaman bakılmıştır. Görülmüştür ki, bu uç değerlerden girişte yüksek olan çıkışta hemen daima düşmüş veya düşük olan değer çıkışta azalmıştır. Söz konusu değerlere hastaların çıkış zamanındaki değerler baz alınıp bakıldığında da bir şey değişmediği görülmüştür. Burada kural normal dağılımda bir uç değer alındığında ondan sonra alınacak değerin çok büyük bir olasılıkla dağılımın ortalamasına daha yakın olmasıdır.

Esasında plasebo etkisi dediğimiz olayın önemli bir bölümü de doğrudan doğruya ortalamaya regresyondur. Diyelim ki, tansiyonu yüksek bir grup hasta aldık ve bunların yarısına tansiyon düşürücü ilaç, diğer yarısına da plasebo verdik. Bu durumda, gerçek ilacı alanlar yanında plasebo alan grupta da, özellikle tansiyonu en yüksek olanların arasında, plasebo sonrası tansiyonu düşmüş buluruz. Bunu nedeni de ortalamaya regresyondur. İlaç çalışmalarında ortalamaya regresyon etkisini azaltmanın en kestirme yolu denekleri olabildiğince uç değerler taşımayanlar arasından seçmektir.

Banu Çakır: Ekstrem değerlerden kaçınmak gerektiğini söylemiştik. Ancak klinikte yapılan bir hata var. Örneğin, kan basıncı ölçerken üç ölçüm alıyoruz ve üç ölçümün ortalamasını alıyoruz. Hastayı ne kadar dinlendirirsek dinlendirelim ilk ölçüm mutlaka yüksek çıkacaktır. Beyaz önlük etkisi nedeniyle ikinci ve üçüncü ölçümün de ortalamasını aldığımızda kişinin gerçek değerini ölçebiliriz. Burada ekstremlerden kaçınmaktan kastedilen ilk ve en yüksek değeri kabul etmemektir. Ancak, alınan değerlerden biri çok yüksek çıkarsa o değer atılmamalıdır. Çünkü o kişinin değeri gerçekten yüksekse ve yüksek değer atılırsa o zaman çok ciddi bir *bias* yapılmış olur. Veri setinden ekstremleri atarken çok hassas davranmalıyız.

III. Karıştırıcılık

Karıştırıcı yanlışlığı veya karıştırıcılık (*confounding*) iki değişken arasında var olduğu gözlenen anlamlı bir ilişkinin esasında üçüncü bir değişken (karıştırıcı veya *confounder*) nedeniyle ortaya çıkmasıdır. Örneğin ağır SLE'li hastalarda hipertansiyon sık görülür. Bunun en önemli nedeni ise yine böyle hastalarda sıkça kullanılan steroidlerdir.

Karıştırıcılık yanlılığını diğer iki belli başlı yanlılık türünden (bilgi edinme ve seçme türleri) ayıran bir nokta, eğer böyle bir yanlılık olduğu kuşkumuz doğarsa, deneyin sonunda, sonuçlar değerlendirilirken bir takım istatistik yöntemlerle çalışmayı bu tür yanlılıktan oldukça başarıyla arındırabilmemizdir. Halbuki bilgi edinme ve seçme türü yanlılıklar kalıcıdırlar.

KATKILAR

Bir Katılımcı: Bir hastalık seçip de onda bir özellik aradığınız zaman (örneğin romatoid artritte ekstrapiramidal rijidite veya sklerodermada kardiyak tutulum ya da romatoid artritte primer hipertansiyon sıklığı) genellikle yüksek bulunuyor ve pratikte gerçekten karşınıza bu kadar sık çıkmıyor. Bu sizin anlattıklarınızdan hangi *bias*'a giriyor, *selection bias*'a mı?

Hasan Yazıcı: Verdiğiniz örnek bir seçme yanlılığı üst başlılığı altında Berkson (komorbidite) türünde bir yanlılıktır. Önemli olan kontrol gruplarının olabildiğince (ama aşırı eşleştirme de yapmadan) varsayımı çürütecek yönde alınmasıdır.

Banu Çakır: *Regression to the mean*'le ilgili bir ekleme yapmak istiyorum. Aslında işin bir biyolojik, bir de yöntem sorunu var; her ne kadar istatistik olarak tekrarlamalarda ortalamaya yaklaşma söz konusu olsa da, biyolojik olarak biliyoruz ki bir takım değişkenlerin seyredişi çok yüksek parametrelerde olabiliyor. Uzun boy gibi bazı genlerin uzun, bazılarının daha kısa boylularda olması gibi. Ve eğer biyolojik olarak sizin değişkeninizin geni, oynama miktarı fazlaysa, sapma değeri sizin kullandığınız yöntemle yakalanma olasılığını da arttırıyor çünkü sizin kullandığınız yöntemin de bir duyarlılığı var. Yani bir biyolojik değişkenin ne kadar farklı sınırlarda oynadığı önemlidir. Bazı değişkenler yüksek olabiliyor, örneğin CRP. Bazı insanlarda CRP biyolojik nedenlerle normalin üst sınırlarında olabilir; bu insanlarda tamamen biyolojik nedenlerle sapma miktarı fazla görünür ve sizin yönteminizin bunu yakalama ihtimali fazladır. İlaç çalışmalarındaki sapma kısmen buna bağlıdır ama biyolojik bir temeli de vardır, istatistik tekrarlanabilirlik dışında.

Hasan Yazıcı: Ayrıca ortalamaya regresyonun olabilmesi için, ölçtüğümüz şeyin normal dağılması gereklidir. Gen çalışmalarında ölçülenler, örneğin bir enzimin *output*'u ya da tek bir gen ekspresyonu, normal dağılım göstermeyebilir. Ölçtüklerimiz boy, kilo, sedimantasyon hızı gibi multifaktöryel değerlerdir. Onlarda normal dağılım olması doğal, normal dağılmıyorsa *regression to mean* olmaz.

Haner Direskeneli: Sizin verdiğiniz *bias* örneklerinde mono ve kombine terapilerle ilgili bir yaklaşım vardı. Eğer dirençli hastaya kombine terapi verirsek pleseboyla karşılaştırmalı çalışmada bu nasıl fark yaratır? O zaman durumu daha hafif olan hastaya vereceğiz. Çalışma dizaynını hafif hastada monoterapi versus kombine terapi yapacağız ve büyük ihtimalle de kombine terapinin farkı hasta hafif olduğu için çıkmayacaktır. Gerçek hayatta da biz genelde monoterapiye dirençli hastaya kombine terapi veriyoruz. Yani bu anlamda dirençli hastaya vermek klinik anlamlılık açısından ve bizim modellememiz açısından yine de doğru, değil mi?

Hasan Yazıcı: Çok doğru ama her zaman randomize kontrollü çalışma yapmak zorunda da değiliz. Randomize kontrollü klinik çalışma iyi bir yöntem ama sadece bir yöntem, bir enstrüman değil mi? Bu enstrümanın bir takım kısıtla-

maları, oyun kuralları mevcuttur. Baştan girme koşulları eşit olacaktır. Penisilinin veya insülinin iyi geldiği randomize kontrollü çalışmayla gösterilmedi. Randomize klinik çalışmayı yanlış yerde kullanmayalım.

Bir Katılımcı: Şöyle bir planlama çok daha uygun olabilir: A ilacına dirençliyse, ya A'ya devam ediyoruz ya kombine veriyoruz. A'ya devam etmek yerine B ilacına döneriz; böylece de sizin söylediğin sorundan da kurtulmuş oluruz. Mesela metotreksata dirençli hastaya kombine tedavi verecek iken metotreksatı, leflunomidle değiştirebiliriz. O zaman monoterapi mi yoksa kombine terapinin mi daha etkili olduğu sorusunun yanıtını almış oluruz.

Melike Melikoğlu: Ben bir şey sormak istiyorum. Hem *regression to the mean* hem de plasebo efektini aşmanın yöntemi zaten kontrollü çalışma yapmak değil mi? Yani *bias*'tan kontrollü olarak kurtulmanın yolu bu değil mi?

Hasan Yazıcı: O zaman ekstrem koşullar alırsak çalışmanın sonunda etkisiz gösterme tehlikesi doğar. *Regression to mean*'den dolayı etkisizlik görülebilir, ortalamadan sapma hatası olabilir, çalışmanın gücü yetmeyebilir.

Melike Melikoğlu: Tam tersi olması lazım...

Hasan Yazıcı: Tamam ama Tip II hata çıkabilir.

Melike Melikoğlu: Ama çıkmamışsa.....

Hasan Yazıcı: Güç konusunu bir kenara bırakırsak dediğiniz doğru ama unutulan bir konudur bu. Yapılan 4000, 5000, 7000 hastalık çalışmaların altında bir miktar da bu yatıyor.

KAYNAKLAR:

1. Yazıcı H, van der Linden S. The case control study and its potential biases. Clin Exp Rheumatol 1996; 14: 3558.
2. Medical Epidemiology Ed: RS Greenberg et al. Lange Medical Books, 4th ed. 2004.
3. Biomedical Bestiary: An Epidemiological Guide to Flaws and Fallacies in Medical Literature Ed: M Michael et al. Little Brown and Company 1984.

EPİDEMİYOLOJİK ÇALIŞMA YÖNTEMLERİ

M. Tevfik Dorak

Doktor, Newcastle University, School of Clinical Medical Sciences (Child Health), Paediatric and Lifecourse Epidemiology Research Group, Newcastle, UK (http://www.dorak.info/epi)

Bu konuşmada epidemiyolojik çalışmalarda kullanılan yöntemlerden bahsedeceğim. Diğer konuşmacılar klinik ilaç denemelerinden bahsedecektir, ancak ben daha çok temel bilimlerde kullanılan yöntemleri özetleyeceğim. Deneysel çalışmalar bu konuşma kapsamının dışında kalacaktır, yalnızca gözlemsel çalışmaları gözden geçireceğiz.

Epidemiyolojik çalışma yöntemlerini şu ana gruplara ayırabiliriz:

I. Deneysel çalışmalar (klinik ilaç denemeleri de bu gruba girer)

II. Gözlemsel çalışmalar:

a) Tanımlayıcı

b) Analitik

- **i.** Olgu kontrol çalışmaları
- **ii.** Kohort çalışmaları
- **iii.** Ekolojik çalışmalar
- **iv.** *Cross-sectional* çalışmalar

Detaylı olarak en çok kullanılan olgu kontrol ve kohort çalışmalarına değineceğim.

Cross-sectional çalışma ne ileri ne geriye dönük olan, yalnızca belli bir zaman kesitindeki durumu belirleyen ve bu gözleme dayanarak çesitli çıkarımların yapılabildiği bir çalışma türüdür. Ekolojik çalışmalarda ise kişiler değil topluluklar incelenir ve grup düzeyinde sonuçlara varılır. Örneğin, Türkiye çapında hava kirliliği düzeyindeki değişiklikler ile kanser olguları sayısındaki değişiklikler gibi. Bulunan ilişki her zaman kişisel düzeyde geçerli olmayabilir, bunu unutmamak gerekir. Analitik çalışmalar genellikle belli bir hastalığın risk faktörlerini araştırmak için kullanılır. Bu faktörler genetik veya çevresel olabilir. Deneysel çalışmalar daha çok ilaç veya diğer tedaviler üzerine yoğunlaşır.

Analitik çalışmalar *cross-sectional* çalışmaların aksine, risk faktörü ile hastalık ilişkisinde zamanlama bakımından da fikir verebilir. Özellikle kohort çalışmalarında bu ilişkinin zamanlaması tam olarak belirlenebilir. Kohort çalışmaları ileriye yönelik, bazen onlarca yıl sürebilen çalışmalardır. Başlangıçta kimlerin risk faktörlerine maruz kalacağı bile bilinmeyebilir. Örneğin tüm şehir halkı çalışmaya alınıp yıllarca izlenebilir. Japonya'da atom bombası atılmasını takiben hayatta kalan tüm halk takibe alınmış ve maruz kaldıkları atom bombası dozuyla kanser, fertilite ve daha başka hastalıklarla ilişkisi araştırılmıştır

(çalışma halen devam etmektedir). Bir başka örnek de klasik Framingham çalışmalarıdır. Bu çalışmada Framingham şehrinin tamamı çalışmaya alınmış ve yıllarca izlenmeyi takiben çeşitli etmenlerle daha çok kalp ve dolaşım sistemi hastalıkları arasındaki ilişkiler ortaya çıkarılmıştır. Amerika'da süregelen beslenme ve hastalık ilişkilerini araştıran kohort çalışmalarında katılımcılardan kan örnekleri alınıp saklanmış ve yıllarca sonra geriye dönülüp, örneğin serum demir ve çinko düzeyleri gibi ölçümler için kullanılıp, kanser riski üzerine etkilerine bakılabilmiştir.

Bu çalışma yönteminde epidemiyolojide çok önemli olan nedensellik konusunda büyük sorunlardan birisi olan tersine neden olma diyebileceğimiz *reverse causation* olayından da kaçınılmış olur. Bu olay neyin neye neden olduğunun çok açık olmaması ile eşdeğerdir. Örneğin, kanser hastalarında adrenal steroid hormonların yüksek bulunması veya kolesterolün düşük bulunması mutlaka steroid yüksekliği veya kolesterol düşüklüğünün kanser riskini artırdığı anlamına gelmez. Kanserin kendisi bu değişikliklere yol açmış olabilir. O yüzden NHANES gibi kohort çalışmalarında, hastaların kanser gelişmeden önce alınıp saklanmış serum örneklerine bakılabildiği için, kohort çalışmaları bu tür ilişkilerin araştırılmasında (*cross-sectional* çalışmalardan) çok daha yararlıdır.

Kohort çalışmalarında henüz hastalık gelişmeden önce sağlıklı kişiler çalışmaya alınıp, *exposure* bakımından ve daha sonra hastalık gelişmesi bakımından izlenir. Örneğin bir sene içinde bütün tıp fakültesi mezunları çalışmaya alınıp bunlardan sigara içenler ile içmeyenler arasında on yıl sonra akciğer hastalıkları arasındaki fark incelenir. İki grupta bulunan insidans oranları karşılaştırılıp rölatif risk bulunur. Diğer ana grup olan olgu kontrol çalışmalarında ise geriye dönülüp, hastalığın oluştuğu bir grup ile sağlıklı bir grubun hangi risk faktörlerine maruz kaldığı (sorgulama, hastane kayıtları, biomarker çalışması gibi yöntemlerle) belirlenir; iki grup arasında o faktörle karşılaşma ve karşılaşmama oranları arasında fark olup olmadığı incelenir. Bu tür çalışma ile rölatif risk değil ancak *odds ratio* hesaplanabilir.

Bir kohort çalışma örneği daha vermek isterim. Bazen etik nedenlerle bazı risk faktörlerini bilinçli olarak incelemek mümkün olmayabilir ama doğal olarak bu risk faktörleriyle karşılaşan bir grubu kullanarak araştırma sorunuza cevap arayabilirsiniz. Benim Alabama Üniversitesi'ndeyken üzerinde çalışma yapabildiğim *multicenter AIDS cohort study* (MACS), yıllar önce ilk AIDS olguları belirmeye başladığında belli başlı Amerikan şehirlerinde yaşayan eşcinsel erkeklerin kaydedilmiş olduğu bir kohort çalışmasıdır. Tabii ki bu kişilere "şimdi gidin de yapabildiğiniz kadar eşcinsel ilişkide bulunun, biz de yıllar sonra sizin üzerinde AIDS çalışmaları yapalım," denmedi. Hatta HIV kapmamaları için ne yapmaları gerektiğine dair eğitim ve hatta ücretsiz olarak kondom da verildi. Ama yine de belli sayıda HIV infeksiyonu gelişti ve AIDS'e ilerledi. Bu doğal gelişim, hangi tür cinsel aktivitenin, hangi yaş grubunun ve daha başka faktörlerin HIV ile infekte olma riskine etki yaptığını ortaya çıkardı. Bu çalışma halen devam etmektedir ve özellikle genetik çalışmalar için çok yararlı bir kohorttur.

Kohort çalışmaları çok uzun süren ve maliyeti çok yüksek çalışmalardır. Peki niye kohort çalışması yapmak isteriz? Bu tür çalışmalar ile gerçek toplumsal hastalık risk oranları belirlenebilir. Olgu kontrol çalışmalarında böyle bir belirleme yapılamaz çünkü zaten hastalığı edinmiş kişilerle başlarsınız. Kohort çalışmasında ise hiçbir hastalık işareti olmayan kimselerle başlayıp doğal olarak ne kadarında hastalık geliştiğine bakılır. Üstelik belli bir risk faktörünün bu orana nasıl etki ettiği araştırılabilir. Kohort çalışmalarında ayrıca risk faktörü ile hastalık arasındaki zamanlama ilişkisi de belirlenebilir. İlk bakışta olgu kontrol çalışmasında da bunun mümkün olacağı düşünülebilir ama bu her zaman için doğru değildir. Sorgulama yöntemiyle geçmişteki olayları araştırıyorsanız, çok önemli *bias*lardan birisi olan hatırlama yanlılığıyla karşılaşacaksınız demektir. Örneğin, lösemili çocukların annelerine gebelikte ilaç kullanıp kullanmadığını sorduğumuzu düşünelim. Pek çoğu mutlaka bir ilaç kullandığını hatırlayacaktır. Ne kadar zararsız bir şey olsa da hatırlayacak ve onu löseminin sebebi olarak gördüğü için mutlaka zararlı bir şey olarak bildirecektir. Buna karşılık, sağlıklı çocuğu olan anneler zararlı bir şey kullanmış olsa bile, çoktan unutmuş olacak ve hatırlaması için de bir sebebi olmadığından bunu bildirmeyecektir. Sonuçta zararlı ilaç kullanımında bir farklılık olmasa bile yapay olarak bir fark bulunacaktır.

İşte bu duruma, yani gerçekten farklı bir sonuç bulmamıza neden olan etmenlere topluca *bias* diyoruz. Bu etmenler sonuçların sapmasına ve gerçekten farklı çıkmasına neden olur. Bu tür *bias*'ların her zaman olmayan bir ilişkiyi varmış gibi göstereceğini düşünmemek gerekir. Bazen tersi olur ve gerçekte olan bir ilişkiyi kaçırmaya da neden olabilir. Kohort çalışmalarında *bias* olasılığı daha düşüktür ama kesinlikle sıfır değildir.

Kohort çalışmaları pek çok yararı olmakla birlikte her zaman yapılması mümkün olmayan bir çalışma türüdür. Öncelikle çok zaman alıcı, maliyeti yüksek olacağını unutmamak gerekir. Bunun yanı sıra seyrek görülen hastalıkları çalışmak için gerçekçi bir yaklaşım değildir. Benzer şekilde latent dönemi uzun olan hastalıklarda da yararlı olmayacaktır. Bu tür hastalıklarda olgu kontrol çalışmaları bütün dezavantajlarına rağmen daha iyi bir seçimdir.

Kohort çalışmalarında akılda tutulması gereken bir diğer husus da kohort etkisi denen durumdur. Çok uzun süren çalışmalarda bazen kohortun başlangıcında çalışmaya dahil edilenler ile sonradan katılanlar arasında fark olabilir. Örneğin BRCA1 geninin meme kanserine yol açma riski (penetransı) 1930 ve 1940'lı yıllara göre bugün daha yüksektir. Yine uzun sürmesi nedeniyle kohort çalışmalarında katılımcılardan ayrılanlar olacaktır. Eğer bu ayrılmalar seçici olarak gerçekleştiyse alınacak sonuçları saptırabilir. Örneğin, hastalık geliştikten sonra ağır hastalığı olanlar bu çalışmalardan daha seçici olarak ayrılmaktadır. Eğer tüm şehir dahil edildiyse zamanla belli bir etnik gruba dahil olanlar seçici olarak o şehirden ayrılmış olabilirler. Bu gibi durumları gözardı etmemeli ve katılım ve ayrılmaları sürekli gözden geçirmeliyiz.

Kohort çalışmaları ileriye yönelik çalışmalardır, ama bazen geriye dönüp bakmanız ya da daha eski bir tarihten itibaren çalışmayı başlatmanız mümkündür. Örneğin, belli bir hastalıkla ilgili çalışacaksınız ama hasta sayınızı artırmak istiyorsunuz. Eğer geçmiş bir zaman diliminde tanı konmuş bütün hastalara ulaşabiliyor ve hepsini çalışmaya dahil edebiliyorsanız (örneğin hiç ölüm veya göç olmadıysa), özellikle genetik çalışmalar için, kohort çalışmanız daha önceki bir tarihten itibaren başlatılabilir. Genetik çalışmalarda *exposure* doğumdan beri sahip olunan genetik faktörler olduğundan ve hastalıkla ilgili gelişmeler hastaların hastane veya doktor kayıtlarında bulunacağı için geriye dönük bir kohort oluşturmak mümkün olmaktadır. Tabii ki bu her hastalık veya her çalışma için mümkün olmayabilir.

Hasan Yazıcı: Kohort çalışmalarında hasta yılı kullanımını ben doğru bulmuyorum.

Tevfik Dorak: Haklısınız. Bazen 100 hastayı bir yıl izlemekle, 10 hastayı 10 yıl izlemek aynı sonucu vermeyebilir. Bazı durumlar bir sene sonra oluşabileceği gibi bazıları için dokuz sene beklemek gerekebilir. Böyle durumlarda hasta yılını kullanmak doğru olmayacaktır.

John P. A. Ioannidis: Kohort ile olgu kontrol çalışmaları arasında kalan bir de eşleştirilmiş kohort çalışması var. Eğer olgu sayısı yetersizse, bu küçük sayıyı ve kontrollerin tamamını analize dahil etmek gerekmeyebilir. Bu durumda, bütün olgular ve eşleştirilmiş ve eşit veya biraz daha yüksek sayıda kontrol dahil edilir ve kohortun kalan kısmı bu çalışma için kullanılmaz. Tabii ki bütün kohort, başka çalışmalarda başlangıçta aklınıza gelmemiş sorulara cevap vermek için yine de kullanılabilir.

Banu Çakır: Başka hibrid tasarımlar da var. Bunlardan bir tanesi *case-cohort*'tur. Tüm kohortu incelemek yerine kohort grubunun bir *baseline* özelliği alınarak, onun içinden randomize, temsil edici bir alt grup seçiliyor ve o kohortun, o küçük alt grubu yıllar boyunca takip ediliyor. Bu altgrubun içinde yıllar boyunca çıkan farklı vakaları (ki bu vakalar diyabetliler, hipertansiyonlular, koroner kalp hastaları olabilir), vaka olarak alıp kohortun o küçük alt grubunu da bir kohort grubu olarak, *efficiency*'yi artırabilmek amacıyla, değişik *case-cohort*'larda aynı kohortu kullanabiliyorlar. Kohortların her zaman prospektif, vaka kontrollerinin de her zaman retrospektif olacağı düşünülüyor. Oysa öyle olmadığını gördük. Bazı vaka kontroller var ki, yuvalandırılmış vaka kontrolü adını veriyoruz. Bunlarda da mesela prospektif bir komponent var. Retrospektif kohort vaka kontrolü ile çok sık karışıyor. Onun için retrospektif ve prospektif kelimelerini çalışmanın tasarımıyla özdeşleştirmemek gerekir. Prospektif ve retrospektif aslında çalışmada *outcome*'dan önce ya da sonra gelmenizle ilgili bir şeydir. *Outcome* oluşmadan önce araştırmacı olarak devreye giriyorsanız, o zaman prospektif, *outcome* oluştuktan sonra devreye

giriyorsanız retrospektiftir. Prospektif ve retrospektif terimlerini çalışma tasarımından farklı olarak ayrı ayrı kullanmak gerekir.

Hasan Yazıcı: Bir de retrospektif çalışmaların sadece klinik çalışmalar olduğu söyleniyor ki, bu tümüyle yanlıştır. Laboratuvar çalışmalarının %100'ü retrospektiftir.

Tevfik Dorak: Bu katkıları özetlemek istiyorum. *Matched-cohort* veya *case-cohort* denilen tasarım ilginçtir ve yaygın olarak kullanılmaktadır. Sonuç olarak, ben de yayınlarımdan birinde kullandım. Onbin kişilik bir kohortta yalnızca 150 kişi beklenen hastalık veya başka bir duruma sahipse kalanların tamamını kontrol olarak kullanmak çok yararlı olmadığı gibi, maliyet açısından zararlı da olabilir. Bu durumda 150 olgu ve örneğin bire üç oranında eşleştirilmiş kontrol kullanarak bir *case-cohort* çalışması yapabilirsiniz. Benim Zambia'da HIV'nin eşler arasında heteroseksüel geçişi üzerine yayınladığım çalışmada bu yöntem kullanılmıştır. Yüzlerce çift çalışmaya alınmıştı ama sadece 100 kadarında HIV geçişi olmuştu. Kohortun tamamını HLA tiplendirmesi için kullanmak israf olacağı için geçiş görülen bütün çiftler ve onlara yaş ve diğer özellikler bakımından benzeyen iki misli kontrol çift, yani HIV geçişi görülmeyen çiftler kullanılarak güzel bir çalışma yapmıştık.

Ne tür çalışma yaparsak yapalım sonuçta gördüğümüz etkiyi sayısal olarak ifade eden bir parametre kullanmalıyız. Yalnızca p değeri vermek çok yetersiz fakat çok fazla gördüğümüz bir durumdur. Bu parametreler çok sayıda ve en çok kullanılanlar rölatif risk ile *odds ratio*'dur, fakat diğerleri de mevcuttur. Bunların tam listesini ve hesaplamak için web üzerinde bir hesap makinesini "*evidence-based medicine toolbox*" web sitesinde (http://www.cebm.net) bulabilirsiniz. Ben özellikle mutlak risk ile rölatif risk arasındaki farka dikkatinizi çekmek isterim. Özellikle klinik ilaç denemelerinde, mutlak riskin daha önemli olduğunu, rölatif risk ne kadar yüksek olursa olsun, mutlak riskte anlamlı bir değişikliğe neden olmuyorsa fazla bir değeri olmadığını vurgulamak istiyorum. Hiçbir epidemiyolojik çalışma tamamen hatasız olmayacaktır. Amaç hataları en aza indirmektir. Bu çalışmalarda elde edeceğimiz sonuçları etkileyecek en önemli faktörler *bias, confounding* ve rasgele hatadır. Bunun yanı sıra etki değişimi ve daha önce sözünü ettiğim *reverse causation*'ı da unutmamak gerekir.

Rasgele hata dediğimiz şans bulgularıdır. Her zaman için tamamen şansa bağlı sonuçlar elde etmeniz mümkündür. Ne kadar beklediğimiz gibi de olsa, her sonuç aldığımızda mutlaka gerçek olduğundan emin olmaya çalışmalıyız. Şans faktörünü dışlamanın başlıca yolları, çok yüksek sayıda olgu ve kontrol kullanmanız ve sonuçlarımızı ikinci bir çalışmada doğrulayabilmemizdir. Aksi halde, aldığınız sonuçlara çok ciddi gözle bakılmayacaktır.

Bias için yapabileceğiniz tek şey dikkatli olmaktır. Sayıyı arttırmak veya istatistiksel yöntemler *bias* sorunu ile başa çıkamaz. *Confounding* aldığımız sonuçların aslında, bir başka ilişkili faktöre bağlı olmasıdır; örneğin, sigara ilişkisinin karşımıza alkol ile akciğer kanseri arasında bir ilişki varmış gibi çıkması. Bu ilişki tamamen alkoliklerin aynı zamanda çok sigara içmesi ile ilgilidir. Bunu önlemenin yolu bilinen *confounding* faktörler ile ilgili veri toplayıp, analiz safhasında bütün bu ilişkileri dikkate almamızdır. Sorun bilinmeyen *confounding* faktörlere bağlı olabilir, o takdirde tek şansınız randomizasyon yöntemidir ki, her çalışma için geçerli olmayabilir. Kohort çalışmaları, bu tür yan etkileri nispeten az olan yöntemdir ancak yine de sorunsuz bir yöntem değildir. Önemli olan klasik epidemiyolojik çalışma prensiplerine bağlı kalmak, mümkün olduğu kadar yüksek olgu ve kontrol sayılarını amaçlamak ve her çalışmayı bağımsız bir grupta tekrarlamaktır. Bu konuşmada kullandığım sunumu internette bulabilirsiniz (http://www.dorak.info/epi).

META ANALİZ

John P. A. Ioannidis

Doktor, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine; Deprtment of Medicine, Tufts University School of Medicine

Meta analiz, birbirine kabaca benzeyen çalışmalardan elde edilen verileri bir araya getirmek için kullanılan bir istatistik yöntemdir. Kimi araştırmacıya göre yöntem, çalışmalarda gözlenen etkilerin de kabaca benzer olması halinde bir anlam taşır. Ayrıca bu çok gerçekçi değildir. Biliyoruz ki elimizde sadece iki çalışma bile olsa, birbirine ne kadar benzer olursa olsun, bir ihtimal pek çok açıdan da farklı olacaklardır. Bu nedenle daha geniş bir açıdan baktığımızda meta analiz, birbirine benzeyen ya da benzemeyen ve birbiri arasında sonuçlar açısından uyum da göstermeyebilecek çalışmalardan, titiz matematik yöntemlerle sistematik olarak kalitatif bilgi sentez etme işidir.

Günümüzde çok çeşitli meta analiz uygulamaları vardır. Herhangi bir hipotezi irdeleyen iki ya da daha fazla çalışma varsa, pratik olarak meta analiz yapmak mümkündür. Herkesin tahmin edebileceği gibi bu yöntem değişik alanlarda uygulanmıştır. Spesifik tedavilerin etkinliğini araştırmak üzere yapılmış çok sayıda meta analiz bulunur. Epidemiyolojik çalışmaların da pek çok meta analizi yapılmış, tanısal testlerin doğruluğu konusundaki uygulamaların sayısı giderek artmıştır (örn., referans 1'de bir Yunan-Türk ortak çalışması görülebilir). Ekoloji, psikoloji, eğitim bilimleri ve sosyoloji bu metodun kullanıldığı önemli alanlardır; tıp bilimlerindeki meta analiz uygulamalarında kullanılan matematik metodolojinin bir kısmı bu alanlardan devşirilmiştir.

Çalışmalar neden birleştirilir? İlk akla gelen yanıt gücü arttırmaktır. Eldeki çalışmalar küçük denek grupları ile yapıldıysa, anlamlı ve güvenilir bir sonuç vermeyebilir. Bu nedenle veriyi bir araya getirerek gücü artırmaya çalışırız. Ama bu bir yerde kendini avutmaktır. Çünkü yaptığımız şey sonuç olarak kötü tasarlanmış, yeterince kuvvetli olmayan çalışmaları bir araya getirmektir. Bazen işe yarayabilir, ancak bizi başlangıçta olduğumuzdan daha sorunlu bir noktaya da getirebilir.

Çalışmaları birleştirmenin ikinci ve belki de daha önemli nedeni ise heterojeniteyi değerlendirerek sonuçlar arasındaki farkları anlamaya çalışmaktır. Çalışmalar arasında kalitatif, tasarım ya da sonuç farkları olabilir. Peki tüm tasarım, içerik ve analizler farklı olsa bile aynı konuyu irdeleyen çalışmaların sonuçları birbirinden

ne kadar farklı olabilir? Bence bu tek başına çok ilginç bir konudur. Buradan hareketle bulgularımızı değişik çerçeve ve durumlar için genelleyip genelleyemeyeceğimizi anlayabiliriz. Sonra heterojeniteyi açıklamaya da çalışabiliriz. Bu çok kolay bir iş değildir ve bol miktarda iyi ve kaliteli veri gerektirir. Ayrıca yanlılıkları saptayabilir ve devre dışı bırakabiliriz. Meta analiz bu konuda çok önemli rol oynayabilir. Meta analiz, saptırmaları irdelemek için kullanılabilecek en verimli yöntemlerden biridir. Bütün verileri bilgisayara koymak ve çıktısını alarak sonuçları değerlendirmek kolay gözükmekle birlikte, gerçekçi değildir. Süreç daha karmaşıktır. Meta analiz yaparken geriye dönüp, çalışmaların ne sorunu olduğunu ister tek tek çalışma düzeyinde, ister alan düzeyinde görebilir, o alanda ne tür saptırmalar olduğunu, bu saptırmaların ne kadar önemli olduğunu ve sonuçlar üzerinde ne şekilde etki edebileceğini anlayabilir ve yanlılıklardan neler öğrenebileceğimizi görürüz. Sonuçta kesin bir yanıta ulaşamasak bile en azından geleceğe dönük çalışmaları bu yanlılıklardan kaçınarak tasarlama şansımız olabilir.

Meta analiz sadece yayınlanmış veriyle yapılmak zorunda değildir. Kanıtları iyileştirmek ve araştırma kalitesini artırmak için en güzel fırsatlardan biri, kişisel veri düzeyinde ve uluslararası işbirliği ile yöntemi hayata geçirmektir. Böylece süreç prospektif bir araştırma tasarımına dönüşerek araştırmacılar arasında işbirliğinin iyileştirilmesine, veri ve fikir paylaşımına, ortak bir amaç ve protokoller üzerinde çalışmaya ve değişik spesifik alanlarda araştırma standartlarının artırılmasına yardımcı olabilir.

Son on yıl içinde kanıt düzeylendirilmesinde meta analizler, tıbbi kanıtlar arasında hep en üst sıraya konmuştur. Onların ardından randomize kontrollü çalışmalar, randomize olmayan prospektif kontrollü çalışmalar, retrospektif kontrollü çalışmalar, kontrolsüz ve tarihi kontrollerle yapılmış çalışmalar, olgu raporları ve son olarak uzman görüşleri ve uzmanlar tarafından yazılmış sistematik olmayan derlemeler sıralanır. Ancak tasarım tek başına bir verinin daha üstün olduğunu göstermez, çalışmanın kalitesi de çok önemlidir. Meta analizlerin popülerleştiği, sağlık bilimleri çalışmaları arasında en çok atıfta bulunulan çalışma tipi haline gelmiş olmasından da anlaşılır. Zaman geçtikçe diğer çalışma tiplerine kıyasla atıf sayılarının daha da fazla arttığı görülmektedir.

Çoğunlukla karşınıza çıkacak meta analizlerin önemli bir kısmı yayınlanmış veriyi temel alan çalışmalardır. Bu çalışmalar mevcut veriyi kullanır ve mevcut çalışmaların detaylı ve dikkatli şekilde aranması ve değerlendirilmesiyle yapılır. Ancak bazı durumlarda meta analizde yayınlanmamış veriler de kullanılabilir ya da analizin tamamı yayınlanmamış verilerle de yapılabilir.

Terminoloji kafa karıştırıcı olabilir. "Meta analiz" muhtemelen standart terimdir; bunun yanında "kümülatif meta analiz" diye sık kullanılan başka bir terim de vardır. Kümülatif meta analizde o alanda yapılmış tüm çalışmaların özeti olan bir hesap yerine, özet etkinin gelişimini (ve dolayısıyla heterojeniteyi) ararız. Örneğin her yıl için ayrı özet hesap yapabiliriz.

Sistematik derleme İngiltere'de çok yaygın bir terimdir, ABD'de o kadar yaygın değildir. Kelimenin bire bir karşılığı olan sistematik derleme ve meta analiz yöntemlerini içerir. Ancak ikisini birbirinden ayırmak gerekir. Sistematik bir yaklaşımla, standart kurallara dayanarak bir konu hakkında bilgi toplayabilirsiniz ancak bu kantitatif bir sentez yapabileceğiniz anlamına gelmez.

Havuzlama analizi (*pooled analysis*) terimi de sıkça rastlanan bir terimdir ancak kanımca bu terimi kullanmak çok iyi bir seçim değildir. Bu terimi eleştirenler "havuzla ya da boğul" ifadesini kullanmıştır. Havuzlamak sözü bir yandan sanki heterojenliğine bakmadan veriyi bir araya koymak gibi bir anlam çağrıştırıyor.

İşbirlikli (*collaborative*) analiz de bir meta analiz olabilir. Burada tasarım, idare, analiz gibi evrelerin bir kısmı ya da tamamının gerçekleştirilmesinde değişik gruplardan araştırmacılar görev alır.

Meta regresyon, meta analize dahil edilen veriye regresyon yöntemi uygulanması anlamında kullanılır. Meta regresyona sokulan veri birimleri meta analize giren çalışmalardır. Çalışmalar yerine tek tek hastaların/katılımcıların verisini kullanarak daha iyi iş yapılabilir ve daha güvenilir sonuçlar elde edilebilir.

Tek tek hasta verilerinin meta analizi ya da tek tek katılımcı verilerinin meta analizinde ise,

bir araya getirilen çalışmalardaki grup hesapları yerine her hasta ya da çalışmaya katılmış her katılımcının verisini tek tek bir araya getirerek tümünü tekrar analiz etme avantajı vardır.

Meta analiz yaparken yayınlanmamış veri bulunması ne kadar sorun yaratır? Bu sorunun yanıtı duruma göre değişebilir. Yayınlanmamış veriyi bulmak, muhtemelen mümkün değildir. Gitmiş, kaybolmuş, imha edilmiştir, kimse bilemez. Ancak bazen yayınlanmamış veriyi bulmak, elde etmek mümkün olabilir. O zaman da bu veriyi nasıl kullanacağımızı düşünmek gerekir. Sponsorlardan başka kimsenin görmediği bir veriyi kullanmak sorun olabilir. Ayrıca sonuçları hesaplarken yayınlanmış veriyi dahil ederek ve etmeyerek iki hesap yapıp bunları karşılaştırmak faydalı olabilir.

Tipik bir meta analizde, her çalışma için bir nokta değer ve etrafındaki güven aralığı (genellikle % 95) hesaplanır. Kümülatif meta analiz çerçevesinde, her nokta ve güven aralığı, o çalışmaya kadar yapılmış olan tüm çalışmalardan gelen özet etkiyi gösterir. Meta analizler negatif ya da pozitif sonuçlara varabilirler. Bu aslında yanlış bir adlandırmadır çünkü burada asıl odaklanmamız gereken şey, formel istatistik anlamlılığın varlığı ya da yokluğu şeklinde bir kalitatif dikotom ifadeden ziyade, etki büyüklüğü, belirsizlik ve heterojenliktir. Bir başka yanlış inanış da negatif sonuç vermiş meta analizlerin faydalı olmadığı inancıdır. Eğer güven aralıkları klinik açıdan önemli olabilecek bir etki olmadığını gösteriyor, ve *null* sonuca neden olacak saptırmalar dışlanabiliyorsa aslında negatif meta analiz sonuçları son derece faydalı bile olabilir. Burada net bir yanıt alınmıştır ve halihazırda yanıtlanmış soruyla ilgili yeni çalışmalarla uğraşmayı bırakıp, başka sorulara yönelebiliriz demektir. Klasik yayın saptırmasının (*publication bias*) yanında gecikme saptırması da (*time lag bias*) görebiliriz; bunun anlamı aynı soru hakkında negatif sonuç veren çalışmaların pozitif sonuç veren çalışmalardan daha geç yayınlanmasıdır. Çalışmaların kayıt altına alınmaya başlanmasıyla birlikte yayın saptırmasının azalacağına inanıyorum. Böyle olması yayın saptırmasını huni grafiği (*funnel plot*) gibi düzmece istatistik yöntemlerle tanımlamaya çalışmaktan çok daha iyidir. Ancak gecikme sapması hala bir tehdit olarak varlığını sürdürür ve kanıtın ortaya çıkması sırasında gereğinden daha iyimser etki büyüklükleri hesaplamamıza neden olur. Gerçeğin ortaya çıkmasını beklemek zorunda kalabiliriz ve çok etkili tedaviler konusundaki ilk umutların çoğu kez abartılı gözlemler olduğu ortaya çıkar.

Bilim ve yayınlanan araştırmaların global perspektifini de düşünmek gerekir. Örneğin hepimiz araştırmaların tümünün ya da çoğunun bir şekilde son noktada bir İngilizce dergide yayınlanacağına inanırız. Ancak iş böyle değildir. Örneğin Çince yayınlanan 8000'den fazla dergi mevcuttur ve bunların büyük bir kısmı PubMed'de endeksli değildir. Çince literatürün İngilizce literatürden farklı eğilimleri olduğunu gösteren veriler vardır. Örneğin genetik epidemiyoloji konusunda Çin literatürü hemen daima anlamlı etkiler bildirir ve bu çalışmaların meta analizi muhtemelen son derece yanıltıcı olacaktır.

Son olarak seçici sonuç ve seçici analiz, meta analiz açısından önemli bir tehdit oluşturur. Tüm çalışmalar bildirilse ve veriler elde olsa da analiz ve sonuç bildiriminin seçici yapılması halinde belirgin saptırmayla karşı karşıya kalınır ve etki varmış gibi sonuçlar çıkabilir.

Kanımca meta analizleri, tek çalışmaların formal istatistik anlamlılığa ulaşamaması nedeniyle onları bir araya koyup güç artırarak anlamlılık aramak şeklinde algılarsak, bu son derece yanıltıcı ve zararlı bir yaklaşım olur. Bence meta analizleri bir alandaki kanıtları olabildiğine objektif ve saydam şekilde enine boyuna kavramak şeklinde algılamalıyız. Daha ideal olarak alanın sınırlarını bilmeliyiz. İnsanların yeni çalışmalar yapmasını bekleyip ardından bilgileri (yani parçaları) olaydan sonra toplamak yerine, herkesin ilgisini çeken sorular üzerinde dünya çapında araştırmacı ekipler kurup prospektif meta analiz yapmak üzere güçlerimizi birleştirebiliriz. Örneğin romatoid faktörün genetiği hakkında dünyada çalışan 100 ekip bulunabilir. Her ekip kendi araştırma programı çerçevesinde, istediği şeyi ilk bulan olup yayınlayabilir ve buluşun övgüsünü toplayabilir; ancak kalan 99 ekibin bilgisiyle biz bu bilgiyi sistematik olarak tekrarlayıp geçerliliğini sıkı meta analiz yöntemleriyle sınayabileceğimizi de biliriz. Yani meta analiz kavramı yavaş yavaş uluslararası işbirliği ve prospektif tasarımlara doğru kaymaktadır.

KAYNAKLAR:

1. Karassa FB, Afeltra A, Ambrozic A, Chang DM, De Keyser F, Doria A, Galeazzi M, Hirohata S, Hoffman IE, Inanc M, Massardo L, Mathieu A, Mok CC, Morozzi G, Sanna G, Spindler AJ, Tzioufas AG, Yoshio T, Ioannidis JP. Accuracy of anti-ribosomal P protein antibody testing for the diagnosis of neuropsychiatric systemic lupus erythematosus: an international meta-analysis. Arthritis Rheum 2006;54:312-24.

2. Lau J, Ioannidis JP, Schmid CH. Summing up evidence: one answer is not always enough. Lancet 1998;351:123-7.

3. Ioannidis JP, Gwinn M, Little J, Higgins JP, Bernstein JL, Boffetta P, Bondy M, Bray MS, Brenchley PE, Buffler PA, Casas JP, Chokkalingam A, Danesh J, Smith GD, Dolan S, Duncan R, Gruis NA, Hartge P, Hashibe M, Hunter DJ, Jarvelin MR, Malmer B, Maraganore DM, Newton-Bishop JA, O'Brien TR, Petersen G, Riboli E, Salanti G, Seminara D, Smeeth L, Taioli E, Timpson N, Uitterlinden AG, Vineis P, Wareham N, Winn DM, Zimmern R, Khoury MJ; Human Genome Epidemiology Network and the Network of Investigator Networks. A road map for efficient and reliable human genome epidemiology. Nat Genet 2006;38:3-5.

4. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schunemann HJ, Edejer TT, Varonen H, Vist GE, Williams JW Jr, Zaza S; GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ 2004;328:1490.

5. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. N Engl J Med 1992;327:248-54.

6. Ioannidis JP, Rosenberg PS, Goedert JJ, O'Brien TR; International Meta-analysis of HIV Host Genetics. Meta-analysis of individual participants' data in genetic epidemiology. Am J Epidemiol 2002;156:204-10.

7. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. BMJ 2006; 333:597-600.

8. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. JAMA 2005;294:218-28.

9. Pan Z, Trikalinos TA, Kavvoura FK, Lau J, Ioannidis JP. Local literature bias in genetic epidemiology: an empirical evaluation of the Chinese literature. PLoS Med. 2005;2:e334.

GENETİK ÇALIŞMALARDA YÖNTEM

Ahmet Gül

Profesör Doktor, İstanbul Üniversitesi, İstanbul Tıp Fakültesi, Romatoloji Bilim Dalı

Toplantının araştırmalarla ilgili bölümünün en az iki konuşmasının genetiğe ayrılmış olması, aslında şu anda dinleyicileri oluşturan araştırmacıların önemli bir bölümünün genetik çalışmalar da yapmakta olmasına dayanıyor. Fakat, genetik çalışmalarla ilgili ciddi sorunlar var ve bu sorunlar da klinik ilaç veya epidemiyoloji çalışmalarında olduğu gibi, bilimsel çalışmalarla ilgili temel prensiplerin algılanmasıyla ve uygulanmasıyla ilgili hatalara dayanıyor. O nedenle ben bu oturumun ilk bölümünde genel olarak "biz bu genetik çalışmaları neden yapıyoruz ve bu çalışmaları yaparken hangi yöntemleri kullanıyoruz?" soruları üzerinde duracağım. Benden sonraki konuşmada ise, Dr. Dorak bu konuda sık karşılaştığımız sorunlar ve bu sorunların çözüm yolları ile ilgili bize yardımcı olacaktır.

Sydney Brenner'ın şu sözünü çok seviyorum ve "niçin genetik inceleme yapıyoruz?"un en güzel cevaplarından birisi olarak kabul ediyorum: "Canlılarla ilgili en temel özellik hepsinin gen taşımasıdır. Bu nedenle, bana göre canlı sistemlerle ilgili tüm açıklamaların bu düzeyde, gen düzeyinde olması gerekir". Genetik araştırmalarımızın ana hedefini de bu şekilde tanımlayabiliriz.

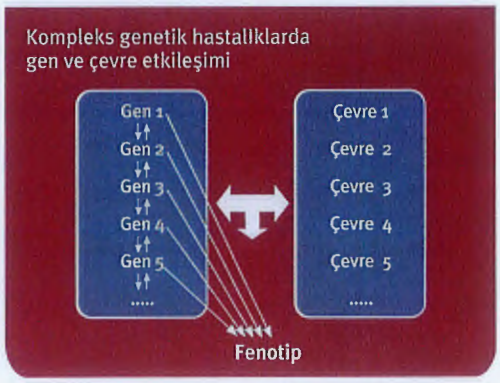
Önce mevcut kaynaklarla başlayacak olursak, bugün insan genomuyla ilgili elimizde ciddi düzeyde bir bilgi olduğunu söylemeliyiz. İnsan Genomu Projesi'nin tamamlanmasıyla, gerçek kromatin yapısında olan DNA'nın tamamını büyük bir doğrulukla bilir hale geldik (www.genome.gov). Yaklaşık 3 milyar baz çiftinden (3 Gb) oluşan insan genomunun, % 1'den biraz fazla bölümü protein kodlamaktadır. Protein kodlayan genlere RNA genleri de eklense, bu dizilerin genomun küçük bir kısmını oluşturduğu görülmektedir. Genomun önemli bir bölümü, aslında tam olarak ne işe yaradığını bilmediğimiz bir yapıda olan ve büyük çoğunluğu tekrarlayıcı veya bir yerden bir yere nakledilmiş (transpoze) dizilerden oluşmaktadır. Bu genom bölgelerinin, herhangi bir işe yaramayan *junk* özellikte diziler olmadığını ve ciddi oranda düzenleyici rolünün olduğunu düşünüyoruz. En azından, özellikle genetik epidemiyoloji açısından, türümüzün geçmişi ile ilgili bilgilerin orada saklı olduğunu biliyoruz. Dolayısıyla, genomun gen kodlamayan bölümünün de mükemmel bir kaynak olduğunu söyleyebiliriz.

Genlerdeki bozukluklar, özellikle tek gen hastalıkları, genetik çalışmaların asıl zeminini oluşturmuştur. Monogenik hastalıklar veya tek gen hastalıkları dediğimiz hastalıklar 10,000 kadar hastalığı içine almaktadır. Monogenik hastalıkların tamamı, aslında bütün hastalıklar içerisinde % 1 gibi küçük bir kısmı oluşturur. Dolayısıyla, çok nadir görülen hastalıklardır ve farklı kalıtım şekilleriyle sonraki nesillere iletilir: otozomal dominant, otozomal resesif, X kromozomuna bağlı, mitokondri kalıtımı, *imprinting*, üçlü nükleotid tekrarı gibi. Sonuç olarak, farklı şekillerde kalıtım gösteren, tek gendeki bozukluklarla meydana gelen ve nadir görülen hastalıklardan söz ediyoruz. Biz bu hastalıkları incelerken, her şeyden önce kalıtım paternini biliyoruz. Örneğin, dominant bir kalıtım modelinde biliyoruz ki, hastalık fenotipi ebeveynlerden birisinin çocuğa nakledeceği gene bağlıdır ve çocuk o geni aldığı anda (genin penetransı oranında) bu hastalığın ortaya çıkması söz konusudur. Bu olasılığın ne kadar olduğunu, gendeki mutasyonun penetransı belirler. Söz konusu gendeki değişikliği (mutasyonu) taşıyor olup da, hastalık fenotipinin görülebilme olasılığına penetrans diyoruz. Otozomal dominant geçiş gösteren hastalıklarda, tam penetran bir mutasyonu taşıyan herkeste hastalık fenotipi görülür; anne ve babadan geçen genlerden birisi sağlam ve sadece birisinde bu mutasyon olsa bile. Otozomal resesif hastalıklarda, yani çekinik geçen hastalıklarda, hastalık fenotipinin ortaya çıkması için, hem anneden hem babadan mutasyonun naklediliyor olması gerekir. Ancak, her iki gende de mutasyon taşınması durumumda hastalık belirtilerinin görülmesi mümkün olur. Eğer siz hastalığın kalıtım paternini biliyorsanız, çocuklarda hangi olasılıkla hastalık oluşabileceğini de tahmin etme şansınız vardır. İşte hastalıklı ailelerde, bu tahmin yöntemlerini kullanarak ve mutasyonlu genlerin kalıtılmasını izleyerek genleri tanımlayabilme şansımız vardır ki, bu yönteme bağlantı analizi diyoruz. Kalıtım paternini ve penetransını bilmek için aslında soy ağaçlarını incelemek yeterlidir. Elinizde birkaç nesli izleyebileceğiniz çok sayıda aile varsa, bu ailelerin incelenmesi yoluyla kalıtım paterni söylenebilir ve penetrans tahmin edilebilir. Bu şekilde Mendel kuralları ile kalıtım gösteren ve penetransı bilinen bir hastalıkta, farklı kromozom bölgelerini gösteren işaretleyicilerle hastalık geninin birlikte kalıtıldığını, yani bir sonraki nesile birlikte geçtiğini gösterme yoluyla bağlantı analizi yapılabilir. Bu analiz gen ve işaretleyici arasındaki uzaklığa ve rekombinasyon olasılığına dayanmaktadır. Eğer sizin seçtiğiniz işaretleyici ve hastalıkla ilişkili gen birbirlerine yakınlarsa, Mendel kurallarına göre birlikte kalıtım olasılıkları yüksek, birbirlerinden ne kadar uzaklarsa (ki olabilecek en uzak ihtimal bu genin bir başka kromozomda olmasıdır) birlikte geçme ihtimalleri o kadar düşüktür. Öyleyse biz birlikte kalıtım olasılığına bakarak, o genin kullandığımız işaretleyiciye yakınlığını belirleyebiliriz. Hastalıktan sorumlu gen ve kullandığımız işaretleyici birbirlerine uzaklarsa, rekombinasyon olasılığı artacak, geçiş sırasında meydana gelecek rekombinasyonlarla (*cross-over*) birlikte kalıtım şansı giderek azalacaktır. Bağlantı analizinde, hasta ve sağlıklı bireylerde gözlenen rekombinasyonlar (rekombinan ve non-rekombinanlar) sayılır. Rekombinasyon, normalde aynı haplotip üzerinde bulunan genlerin mayoz bölünme sırasında, *cross-over* ile karşı kromozoma geçmesidir. Genler birbirlerine yakınlarsa rekombinasyon olasılığı azalır ve aynı haplotipte birlikte kalıtılma olasılıkları artar. Ailelerde non-rekombinan (işaretleyici ve hastalık geni aynı haplotipte olan) ve rekombinan (işaretleyici ve hastalık geni ayrı haplotiplerde bulunan) bireylere bakarak, gözlediğimiz sonuçlar ile gen ve işaretleyici arasında bağlantı olmadığında beklediğimiz % 50 olasılık, logaritmik yöntem ile oranlanır. Böylelikle, hastalık geni ile işaretleyicimiz arasındaki bağlantı olasılığı (*odds* oranı) hesaplanır. Genellikle bağlantının anlamlı kabul edilmesi için, Lod (*Logarithm of odds ratio*) skoru dediğimiz değerin 3'ün üzerinde olması beklenir. Bu ise kabaca gen ve işaretleyicinin birlikte geçme olasılığının, bağlantı olmama olasılığına oranla, 1000 kat yüksek olması anlamına gelir. Bu yöntem Mendel'den beri bilinen genetik prensiplerin uygulamasına dayanmaktadır.

Bağlantı analizi sayesinde geçtiğimiz yüzyılda pek çok monogenik hastalığa neden olan genin yerinin belirlenmesi mümkün olmuştur. Tek gen hastalıkları oldukca seyrek olmalarına karşın, doğal olarak ortaya çıkmış bir deney gibi, ciddi gözlem imkanları sunarlar. Bu gözlemler sayesinde farklı metabolizma ve inflamasyon yollarını aydınlatmak mümkün olabilmiştir. Bu nedenle, monogenik hastalıklarının genlerin fonksiyonlarının tanımlanması ve hastalık patogenezlerinin aydınlatılması açısından önemli

katkılar sağlayan bozukluklar olduğunu düşünebiliriz. Monogenik hastalıkların haritalanması ve genin bulunması kolay da olsa, bu hastalıklarda da mutasyon ile hastalık fenotipi arasındaki ilişki değişiklikler gösterebilir ve genellikle "tek gen=tek tip hastalık" bağlantısı kurulamaz. Dolayısıyla, tek gen hastalıklarında da, kompleks etkileşimler olduğu düşünülmektedir. En güzel örneği, hepimizin bildiği Ailevi Akdeniz Ateşi (AAA) hastalığıdır. AAA bir tek gen hastalığıdır, ama zannettiğimizden çok daha kompleks genetik-fenotip ilişkileri göstermektedir. Kistik fibroz da iyi bir örnek olabilir.

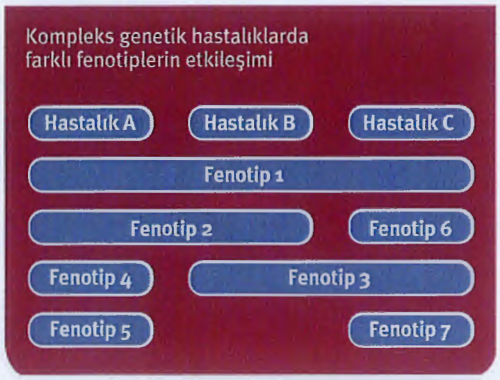
Büyük çoğunluğumuzun uğraşmakta olduğu ve çalışmalarını yoğunlaştırdığı hastalıklar ise kompleks genetik hastalık olarak tanımlanmaktadır. Kompleks hastalıklar dediğimiz zaman, hastalığın patogenezinde çok sayıda gende bulunan değişikliğin katkısıyla oluşan bir genetik yatkınlık ile bazı çevresel faktörlerin etkileşimi sonucunda gelişen hastalık tablosu anlaşılmaktadır. Kompleks hastalıklardaki genetik yatkınlık bir tek gendeki değişiklikten kaynaklanmamaktadır. Hastalığa yatkınlık etkileşimi sonucunda ortaya çıkan durum yol açmaktadır. Buna rağmen, yine Sydney Brenner'dan bir alıntı yapacak olursak, "hepimizin genlerinde bir şekilde geleceğimiz yazılı durumdadır ve görevimiz genlerimizdeki mesajın 'nasıl' etkili olduğunun açıklaması olmalıdır". Bu oldukça zahmetli bir iştir. Özellikle kompleks genetik hastalıklarda, kendi genomumuzdaki çeşitliliğimize bakarak hastalık yatkınlığının nasıl oluşturduğunu açıklamamız gerçekten zordur. İlk olarak, hastalığa yatkınlık oluşması için kaç tane genin etkileştiğini tahmin etmek güçtür. Bu sayı hastalıktan hastalığa ya da hastadan hastaya değişkenlik gösterebilir ve bu genlerin birbirleriyle ve çevreyle etkileşimleri de farklılık gösterebilir (bkz Şekil).



İki gen arasındaki etkileşim aditif olabilir; yani iki polimorfizmi taşıyor olmanızın etkisi, birbirinin toplamı kadar olabilir. Öte yandan iki gen arasındaki etkileşim birbirinin çarpımı kadar olabilir ya da iki gen birbirini potansiyalize edici veya birbirlerinin etkilerini azaltıcı etki gösterebilir. Çevreyle etkileşim sonucunda genin ifadesi (ekspresyonu) artabilir. Bazı çevresel faktörlere maruz kalmak BRCA geninde olduğu gibi genin yapısını veya hastalığın penetransını değiştirebilir. Kompleks hastalıklarda yapılan genetik araş-

tırmalarda, genellikle hastalık fenotipinin oluşması, yatkınlığa neden olan genlerden sadece birisindeki değişiklik ile izah edilmeye çalışılmaktadır. Genellikle proteinlerdeki varyasyonlar (fonksiyonel anlamı varsa) o proteinin işleyişinde, o protein sayesinde organizmada yürüyen işlerde bir değişikliğe, bir fenotip farklılığına neden olabilmektedir. Bunu doğrudan hastalık olarak algılamayalım. Örneğin pıhtılaşmada görevli Faktör V'teki varyasyon pıhtılaşma eğilimini artırabilir veya bir sitokin gen poliformizmi o sitokinin salınım hızını attırabilir. Fakat, bu değişiklikler tek başına bir hastalık durumuna neden olmaz. Bu şekilde birbirinden farklı fenotiplerin birikimiyle bir A hastalığı ortaya çıkabilir (Şekil 2).

Örneğin A hastalığında fenotip 1, fenotip 2, fenotip 4 ve fenotip 5'in katkısı var diyelim. A hastalığı çok sayıda farklı varyasyonun toplam etkisi ile oluşur ve siz fenotip 1'le bir ilişki bulduğunuzda fenotip 1'le A hastalığı arasında bir ilişkiden bahsedebilirsiniz ama bunun doğrudan sebep-sonuç ilişkisi tarzında bir ilişki olmadığı bilinmelidir. Yani, "fenotip 1'e neden olan değişiklik mutlaka A hastalığına neden olur," denemez. Çünkü biliyoruz ki fenotip 1'deki değişiklikler B hastalığıyla ve C hastalığıyla da ilişkili olabilir ve o hastalıkların da patogenezine katkıda bulunabilir.



Dolayısıyla ben bir çalışmayı planlarken aslında bu hastalıkta rol oynayabileceğini düşündüğüm bir fenotipteki değişikliğin A hastalığıyla ilişkili olup olmadığını doğru bir şekilde araştıracak sistemi kurmak zorundayım. Eğer ben bunun yanına kontrol grubu olarak sağlıklı kontrolleri değil de, bir başka hastalığı alıyorsam, bilirim ki bu ikinci bir çalışmadır. Yani araştırdığım fenotipin B hastalığıyla ilişkisini göstermek aslında bir başka çalışmayı oluşturacaktır. Dolayısıyla benim bulduğum ilişki kompleks hastalıklarda asla doğrudan ve tek gen hastalıklarına benzer sebep-sonuç ilişkisi olarak yorumlanmamalıdır.

Bizi diğer insanlardan farklı yapan veya çeşitli hastalıklara yatkın hale getiren ana sebepler nelerdir? İnsan Genomu Projesi tamamlandığı zaman yayınlanan *Nature* dergisinin kapağı çok sayıda insan yüzünden oluşuyordu. Hepsinin saç rengi, göz rengi, yüz görünümü birbirinden farklı olan çok sayıda insan vardı. Bu farklı özelliklerin hepsinin kaynağı genomumuzda bulunan tek

nükleotid polimorfizmleridir (*SNP, single nucleotide polymorphism*). Bunlara basit (*simple*) nükleotid polimorfizmleri demek de mümkün, çünkü polimorfizmlerin hepsi tek nükleotid değişikliğinden kaynaklanmıyor. Bu polimorfizmler hepimizde yaklaşık olarak her 500 ila 1000 nükleotidde bir var olan değişikliklerden oluşmakta ve bu değişiklikler, bizi bir başkasından farklı kılan özelliklere neden olmaktadır. Bu polimorfizmlerin bir kısmı fonksiyonel değişiklikler, yani bir biyolojik değişikliğe neden olabilmektedir. Dolayısıyla, polimorfizmlerin bir kısmı söz konusu protein veya herhangi bir şekilde protein dışında bir başka RNA geninin yapısını ya da ifade özelliklerini etkileyebilen değişikliklerdir ki, o polimorfizmler sayesinde biz bir fenotip değişikliği görebiliyoruz. Bu değişikliklerin bir kısmını, örneğin protein kodlayıcı kısımdaki değişiklikleri anlamak kolaydır. Bazı polimorfizmler ise sessiz değişikliklere neden olabilir; örneğin, GCA, GCG'ye dönüşse bir aminoasit değişikliğine neden olmadığını biliyoruz. Ama örneğin AGA kodonunda G'den A'ya değişiklik, arginini lizine dönüştürüp protein yapısını değiştirebilmektedir. Kodlayıcı bölge polimorfizmlerinin proteinin fonksiyonunu ciddi derecede değiştirmesi şart değildir. Ama, küçük fonksiyonel değişiklikler örneğin benim saçımı kıvırcık, sizinkini düz yapabilir ve bireysel farklılıkları oluşturur.

Hastalıklara yatkınlığa neden olan polimorfizmler sadece kodlayıcı bölgeye özgü değişiklikler değildir. Organizmada proteinlerin sentezlenme yoluna bakacak olursak, sentez uyarısının gelmesinden, DNA düzeyinde transkripsiyon faktörlerinin bağlanmasına, DNA'dan RNA sentezine, RNA'dan protein sentezine, RNA'nın stabilitesine ve ortamda kalma süresine, protein sentezlendikten sonra posttranslasyonel modifikasyonların meydana gelmesine kadar çok sayıda basamak vardır ve bu basamakların her birini etkileyebilecek değişiklikler söz konusu olabilir. Aradığımız gen düzeyindeki değişiklik, bu ana metabolik yolda sadece küçük bir basamağı oluşturabilir. Bazı örnekler vererek anlatmaya çalışacağım. Örneğin, Faktör V Leiden mutasyonunu iyi kötü hepimiz biliyoruz ve günlük klinik uygulamamızda da kullanıyoruz. Faktör V Leiden bir tek nokta mutasyonu (veya polimorfizmi demek daha doğru) ve Faktör V'in protein dizisinde 506. pozisyondaki argininin glutamin ile değişmesine neden olmaktadır. Faktör V proteininin 506. pozisyondaki aminoasit arginin iken aktive olmuş protein C ile parçalanması çok daha kolay iken, glutamin değişikliği sonrasında yıkım zorlaşmakta ve aktive protein C'ye karşı direnç oluşmaktadır. Daha zor yıkılma ve bu sayede de ortaya çıkan direnç, daha uzun süre aktive faktör V proteinine sahip olmanıza ve sonuçta tromboza yatkınlık hali oluşmasına neden olmaktadır. Bununla beraber, fonksiyonel bir değişiklik için aminoasit kodlayıcı bölgede bir değişiklik olması da şart değildir. Örneğin, promotor bölge dediğimiz, o genin sentezlenmesini kontrol eden bölgedeki ya da genin 3' ucundaki değişiklikler, bu genden RNA sentez hızını değiştirebilir, ki bunlara düzenleyici polimorfizmler diyoruz. Sitokinlerin pek çoğunda promotor bölge polimorfizmlerinin olduğunu biliyoruz ve bu değişiklikler çok sayıda hastalıkta çalışılmış durumdadır. Promotor bölge polimorfizmlerinin ne düzeyde etkili olduğunu tam olarak bilemesek bile, o genin sentez hızında değişiklik yapabileceğini düşünüyoruz. Dolayısıyla, benim A genim bir uyarı aldığında 3 hızında mRNA sentezlerken, bir başka insanın geninin 5 hızında sentezleyebildiğini ve bu sentez farklılıklarının bir fenotip değişikliğine neden olabildiğini varsayıyoruz. Düzenleyici polimorfizmlere en iyi örnek protrombin geni 3' ucunda 20210 G → A mutasyonudur ve bu polimorfizmi de klinik pratikte kullanıyoruz. Bu polimorfizmde aslında kodlayıcı bölgede hiçbir değişiklik olmadığını biliyoruz. Protein kodlamayan 3' ucundaki bir değişiklik, sentezlenen mRNA'nın daha hızlı işlenmesine ve bu sayede daha fazla protein sentezlenmesine ve toplam protrombin miktarının daha fazla olmasına neden olmaktadır. Düzenleyici polimorfizmlerde allellerdeki küçük polimorfizmler, o allelin mRNA sentez hızını da değiştirebilmektedir. Örneğin G alleli taşıyorsanız daha fazla mRNA oluştururken, A aleli taşıyanlar daha az mRNA oluşturabilmekte ve böylelikle protein sentezi değişebilmektedir.

Pekala, ben bu bilgileri araştırmalarımda nasıl kullanabilirim? Kendi uğraşmak istediğim hastalık grubunda, (örn., romatoid artrit, sistemik lupus eritematozus veya Behçet hastalığı) bir genetik çalışma planlayayım ve yapayım mı? İlk cevaplamamız gereken soru "Bu hastalıkta genetik çalışma yapmamızı haklı kılacak veriler var mı?" olmalıdır. Araştırma yapmak istediğimiz hastalıkla ilgili elimizdeki bilgilerin, hastalık

patogenezinde genetik katkının olduğunu gösterip göstermediğini ve eğer genetik katkı varsa, bu katkının emek ve para ayırmamıza değecek büyüklükte olup olmadığını bilmemiz gerekir. Çünkü genel anlamda genetik katkı dediğimiz zaman, bazı insanların daha fazla trafik kazası geçirmesinde bile bir yatkınlık tanımlayabilirsiniz, ama bu katkının büyüklüğü gerçekten sizin yatırım yapmanıza değmelidir.

Genetik katkının varlığı ve büyüklüğü konusunda bilgi veren çalışmalara en iyi örnek aile çalışmalarıdır. Ailelerde akrabalar arasındaki hastalık riskinin toplum riskine oranlanması ile, λ (lambda) dediğimiz değerler elde edilir. Bunu bulmak için örneğin kardeşlerde çalışma yapıyorsanız, hastaların kardeşlerinde aynı hastalığın görülme riskinin toplumda hastalık görülme riskine (toplum prevalansı) oranını hesaplarız ve çıkan sonucu kardeş (*sibling*) değeri olarak veririz. Bunu kardeşler dışında, ebeveynlerde de yapmanız mümkündür. Örneğin yaptığımız çalışmadan biliyoruz ki, ülkemizde Behçet hastalığı için kardeş değeri yaklaşık 11 ile 52 arasında ve ciddi bir genetik yatkınlığın varlığını doğruluyor. Ailevi vakalarınız varsa segregasyon analizi de yapabilirsiniz, yani hastalık ailelerde ne tür bir kalıtım paterni gösteriyor buna bakabilirsiniz. Örneğin, birden çok bireyde Behçet hastalığı olan bir aileyi örnek alalım. Kompleks bir hastalık da olsa, ailevi vakalarda bazı soyağaçlarını bir genetikçiye gösterdiğiniz zaman bunun bir monogenik hastalık gibi davrandığını söyleyebilir; üç nesilde hasta olan bir aile ve çok sayıda hasta var. Ama biliyoruz ki, bu kompleks genetik geçiş gösteren bir Behçet ailesi. Behçet hastalığında en azından bir grup hastada pozitif bir aile öyküsü bulmanız mümkündür. Bu oran %10 ila 18 arasında değişebilir. Eğer biz hastalar arasında farklı risk grupları tanımlayabiliyorsak, yani genetik riskin daha fazla olduğu grupları tanımlayabiliyorsak, genetik çalışmalara o gruptan başlayabiliriz.

Yüksek riskli grupların kullanılmasına en iyi örneklerinden birisi meme kanseri olabilir. Biliyoruz ki, 40 yaşından önce meme kanseri oluşmuş hastaları alırsanız genetik yatkınlığın daha fazla olduğu bir grubu tanımlayıp geni tanımlama şansınızı artırabilirsiniz. Böylece genetik katkının ve genlerdeki penetransın daha yüksek olduğu bir grupta çalışma şansınız olacaktır.

Meme kanseri ile ilişkili BRCA geni bu sayede tanımlanmıştır. Pekiyi, bu yüksek riskli grupta elde edilen sonuçlar genellenebilir mi? Yani bu gruptaki gözlemler bütün hastaları açıklayabilir mi? Bütün hastaları açıklayamayabilir, ama bu genlerin tanımlanması sayesinde, söz konusu hastalıkta, belirli fenotipin oluşmasında kritik rolü olan proteinleri tanımlayabilme şansımız olur. Bu bazen resmin bütününü bize göstermeyebilir; özellikle kompleks hastalıklarda bu resmin tamamını görmek oldukça zordur. Ama, hastalık patogenezinde rol oynayan önemli bir yolu aydınlatacağına şüphe yoktur.

Genetik katkıyı göstermek amacıyla ikizler de kullanılabilir: ikiz çiftlerinde hastalık görülme oranına (konkordans) baktığımız zaman, tek yumurta veya çift yumurta ikizlerindeki konkordans bize genetik katkı konusunda fikir verebilir. Fakat ikiz çalışmalarında, konkordans değerlerine hastalığın toplumdaki prevalansının da katkıda bulunduğunu belirtmek gerekir. Eğer toplumda prevalans düşükse konkordans da bir miktar düşecek, sık görülen hastalıklarda konkordans biraz daha fazla olacaktır. Öyleyse konkordans, birebir genetik katkının oranını göstermede yardımcı değildir. Daha uygun olanı bu verileri kullanarak kalıtımsallık (*heritability*) analizi yapmaktır. Kalıtımsallık analizi konkordans verisiyle beraber genetik katkı ve çevrenin etkileşimini bir model üzerinde inceler ve hastalık fenotipinin hangi oranda genetik faktörlerle açıklanabilir olduğunu modellemeye dayanarak tahmin etmeye çalışır. Bir örnek verecek olursam, romatoid artritte tek yumurta ikizlerinde konkordans yaklaşık % 15 civarındadır. Fakat bir modelleme yaparak kalıtımsallık analizi yaptığınız zaman kalıtımın katkısının % 60–65 civarında olduğunu bulursunuz ki, burada özellikle hastalığın toplumdaki prevalansını ve diğer faktörleri de göz önüne almanız gerekir.

Çalışmak istediğiniz hastalıkta genetik katkıya inanıyorsanız, bir sonraki aşamaya geçebilirsiniz. Bu, "Genetik katkı hangi gen(ler)den kaynaklanmaktadır?", "Hangi genom bölgesini taramalıyım?" ya da "Hangi genleri aday genler olarak seçmeliyim?" olabilir. Bu amaçla iki yol izlenebilir.

Bu yöntemlerden ilki aday gen yaklaşımıdır. Aday gen yöntemi için patogenez konusunda ön bilgiye ihtiyaç vardır. Elinizde yeterli veriniz var-

sa, patogenezde rol oynayabilecek proteinleri tahmin edip, o proteinlerin genlerinden yola çıkabilirsiniz. Dolayısıyla, aday gen yönteminde yola çıktığınız nokta, sağlam bir patogenez bilgisine dayanmalıdır. Bu konuda çok fazla bilginiz yoksa, genomda genel tarama yapabilirsiniz. Bu yöntemde düzenli aralıklarla konmuş işaretleyicilerle tüm genomu ya da genomun bir bölgesini, örneğin sadece bir kromozomu tarayabilirsiniz. Eğer aday genlerle çalışıyorsanız en çok uygulanan yöntem vaka kontrol ilişki çalışmaları yapmak veya aileleri kullanmaktır. Ancak, vaka kontrol ya da aile çalışmalarının farklı amaçlara hizmet ettiğini birazdan söyleyeceğim. Eğer aday gen değil de, genom taraması yapacaksanız, bugüne kadar bunun ideal yolu çok hastalı aileler bulmaktır. Bugün için vaka kontroller kullanılarak tüm genom ilişki (asosiyasyon) çalışmaları ile genom taraması da gündemde olmakla birlikte, henüz bu yöntemin genel kabul gördüğünü söylemek pek kolay değildir. Hem istatistik analizinde çözülmemiş konular vardır, hem de henüz sonuçlar erken düzeydedir. Yine de, ailevi vakaların toplanmasının güç olduğu durumlar için gelecek vaad eden bir yöntem olarak görülmektedir.

Konuşmamın bundan sonraki kısmına temel teşkil etmesi için, biraz ilişki çalışmasının temelinden bahsedeceğim. İlişki çalışmaları en çok yapılan genetik araştırma yöntemini oluşturmaktadır. Bu yöntemde, araştırdığınız fenotipe sahip vakalarınız ve sağlıklı kontrolleriniz vardır. Siz araştırdığınız genetik polimorfizm açısından, vakalar ve kontroller arasında fark olup olmadığını öğrenmeye çalışırsınız. Bu karşılaştırmaları yaparken akılda tutulması gereken çok önemli bir nokta şudur: Aslında ilişki çalışması yapılırken, büyük bir aile çalışması yapılmaktadır. Burada söz konusu olan 50-100 nesilli bir ailedir ve vaka ve kontroller açısından aynı kökenden gelen, uzak bir atada oluşmuş bir mutasyonun bugüne yansıyan etkileri karşılaştırılmaktadır. Oluşan mutasyonun kendisinin (fonksiyonel polimorfizm) ya da komşuluğunda bulunan genetik işaretleyicinin (tek nokta mutasyonu ya da mikrosatellit polimorfizmleri) beraber geçme olasılığı test edilir. Çok eski ortak atada oluşmuş bu değişiklik, araştırdığınız fenotiple (hastalıkla) beraber geçiş gösteriyor ise, siz bunu ortak atadan nesiller sonra da görebilir ve vakalar ile sağlıklı kontroller arasındaki farklılığı tanımlayabilirsiniz. Bu farklılığın gösterilebilmesi için, vakalarla aynı kökenden gelen bir kontrol grubunun seçilmesi şarttır. Çünkü, aynı başlangıç noktasından, aynı kaynaktan gelmeyen bir kontrol seçildiğinde ve özellikle aranan fenotip açısından da bu ayrım yapılmamışsa, kontrol grubunun herhangi bir anlamı kalmaz ve vakalarla kontroller arasındaki farkı gösterebilme şansı yok olabilir. Aday geninde seçtiğiniz değişiklik fonksiyonel bir polimorfizm olabilir, o zaman işiniz gerçekten kolaydır. Faktör V Leiden mutasyonu buna örnek olarak verilebilir. Bir ilişki tanımladığınızda, doğrudan fonksiyonu da tanımlamış olursunuz. Bununla beraber, tanımlanmış fonksiyonel polimorfizmler sayıca azdır ve ilişki çalışmalarında genellikle bağlantı dengesizliğine dayanan bir yol izlenmektedir. Aday genlerde, hastalıkla ilişkili fonksiyonel polimorfizm(ler) ile bağlantı dengesizliği gösterebilecek işaretleyiciler kullanılmaktadır. İlişki çalışmalarında kullanılan polimorfizmler ile araştırılan fenotip arasında bir ilişki tanımlandığı zaman, bu ilişkinin nedenleri arasında fonksiyonel mutasyon, etnik gruplar arasında karışım (*admixture*), toplum katmanlanması veya bağlantı dengesizliği yer alabilir.

Anadolu yüzyıllar boyunca çok sayıda göç almış ve farklı etnik grupların kaynaştığı bir coğrafya olmuştur. Göçler etnik kökenlere göre farklı polimorfizmleri de beraberinde getirir ve iki farklı etnik grubun karışması sonucunda kuvvetli bir bağlantı dengesizliği oluşabilir. Bağlantı dengesizliğinin etkisini kaybetmesi çok uzun zaman alır. Ayrıca, toplumdaki katmanlama da bağlantı dengesizliği farklılığına ve hatalı ilişki sonuçlarına neden olabilmektedir. Toplum katmanlaması ülkemiz için ciddi bir sorun olabilir. Sosyal ya da dini nedenlerle izole gruplar halinde yaşayan ve kendi içinde evlenen topluluklar nedeniyle polimorfizmlerin dağılımı değişiklik gösterebilir. Bu konu için Manchester'da çalıştığım grubun bir örneğini vereceğim. Manchester'ın meşhur bir Çin lokantası var. Öyle bir çalışma planlayalım ki, bu Çin lokantasında yemek yiyen insanlarda yemek çubuğu (*chopstick*) kullanma becerisiyle HLA arasında bir ilişki arayalım. Eğer çalışmayı böyle planlarsanız mutlaka çubuk kullanma becerisiyle HLA arasında kuvvetli bir ilişki bulursunuz, çünkü biliyoruz ki o lokantaya gelen Çin kökenli olanların bu çubuğu kullanma becerisi daha fazla olacak, İngiliz kökenlilerin olmayacak. Çinli ve İngiliz kökenli insanların HLA dağı-

lım farklarına dayanarak da, bir ilişki bulunacaktır. Bu bir toplum katmanlaması örneğidir. Toplumdan kendi içinde evlenen, ayrı yaşayan bir grubu alır ve farkına varmadan bunu incelerseniz yanlışlıkla bir ilişki tanımlayabilirsiniz.

Bağlantı dengesizliğinin tipik bir örneği de HLA ve hemokromatoz ilişkisidir. Hemokromatoz HLA-A3 ile ilişkisi olduğu bilinen bir hastalıktır, ama immünolojik bir hastalık değildir. Genetik incelemeler sonucunda, hastalıkla ilişkili olan genin HLA-A'nın telomerinde yer alan HFE geni olduğu, ama bu gendeki mutasyonun HLA-A3 ile kuvvetli bağlantı dengesizliği gösterdiği anlaşılmıştır. HLA bölgesinin özellikleri nedeniyle, bu bölge kuvvetli bağlantı dengesizliği sonucu genellikle blok halinde geçmekte ve mutasyonun ilk oluştuğu haplotipte HLA-A3 bulunduğundan HFE mutasyonu ile HLA-A3 birlikte nakledilmektedir. Dolayısıyla, tanımladığımız HLA-A3 ilişkisi hastalıkla ilişkili bir mutasyon olarak değil, bağlantı dengesizliği gösterdiği HFE mutasyonu sonucunda oluşmaktadır.

Bir vaka kontrol çalışması yaparken, her şeyden önce araştıracağım fenotipi net bir şekilde belirlemem gereklidir. Sağlıklı kontroller hastalar ile aynı kaynaktan geldiği bilinen ve araştırılan fenotipin olmadığından emin olunan bireyler arasından seçilmelidir. Türkiye'de kontrollerin seçilmesi konusunda sorunlarla karşılaşılması beklenebilir. Toplumumuzun oluşumunda 40'tan fazla etnik grubun katkısının olduğu bilinmektedir ve etnik kökenler açısından bölgesel farklılıklar söz konusu olabilir. Örneğin, İstanbul Tıp Fakültesi'nde bir çalışma yapıyorsanız ve kontrollerinizi hastane personeli arasından seçiyorsanız, kontrollerinizin büyük çoğunluğu Kastamonulu olacaktır. Eğer siz bu çalışmayı Hacettepe'de yapıyorsanız, biliyoruz ki kontrollerinizin büyük çoğunluğu Kırşehirli veya Çankırılı olacaktır. Oysa, araştırdığınız hastalar daha geniş bir kaynaktan size ulaşıyor olabilir. Dolayısıyla burada ciddi bir hata söz konusu olabilir. Vakalarınız ile aynı kaynaktan kontrol almayı beceremiyorsanız, ilişki çalışmalarınızın sonuçları güvenilir olmayacaktır.

İlişki çalışmalarında vaka ve kontrol seçimi ile ilgili sorunlardan kurtulmanın yollarından birisi, hasta ve ebeveynlerini kullanarak geçiş dengesizliği testini (*transmission disequilibrium test-TDT*) kullanmaktır. Bu yöntemde hasta ve ebeveyn üçlülerinin genotip verileri kullanılır. Heterozigot olan ebeveynlerden allellerin çocuğa geçiş olasılıkları değerlendirilir. İki allelden birisinin % 50 olan geçiş olasılığındaki sapma değerlerine bakılır. TDT yönteminin üstünlüğü, aynı anda hem bağlantıyı, hem de ilişkiyi gösterebilmesidir.

Diğer önemli konu, ilişki çalışmalarının güvenilir olabilmesi için mutlaka yeterli istatistik güce sahip olmasının gerekliliğidir. Bu amaçla, çalışmaların planlanması aşamasında güç analizi yapılmalıdır. Güç analizi yapabilmek için aday gendeki tek nükleotid polimorfizmlerinin allel frekanslarının bilinmesinde yarar vardır. Bu amaçla daha önce yapılan çalışmaların sonuçları ya da tek nükleotid polimorfizmlerine ait bilgileri içeren veri tabanları kullanılabilir (*Entrez SNP* ve *HapMap*). Bu veritabanları kullanılarak hedeflenen polimorfizmlerin frekans datası bulunabilir ve allel frekanslarında ne kadar değişikliğin anlamlı kabul edileceği kararlaştırılarak güç analizi yapılabilir. *HapMap* veritabanında Avrupalı beyaz ırk, Japon, Çinli ve Afrikalı gruplara ait allel frekansları bulunmaktadır. Bizim allel frekanslarımızın genellikle Avrupalı beyaz ırkla uyumluluk gösterdiği söylenebilir.

Yeterli sayıda ailevi vaka varsa, nonparametrik yöntemler kullanılarak bağlantı analizi de yapılabilir. Başta da söylediğim gibi, genetik geçiş modeli bilinmediği için, bağlantı analizi kompleks hastalıklar için uygun bir yöntem değildir. Fakat, nonparametrik yöntemler kullanılarak, hasta kardeşlerin allel paylaşımı bilgileri ile genom taramaları yapılabilir.

Bağlantı analizinin sonuçları bize hastalıkla ilişkili geni içeren geniş bir genom bölgesini işaret eder ama doğrudan geni göstermez. Bağlantı gösterilen genom bölgesinde onlarca, hatta yüzlerce farklı gen olabilir. Dolayısıyla bir sonraki basamak, bağlantı saptanan bölgedeki hastalıkla ilişkili genleri tanımlamak olacaktır. Oysa, ilişki çalışmalarında aday bir geniniz vardır ve bu gende aradığınız fenotiple ilişkili olduğu düşünülen polimorfizm aranmaktadır. Bu iki araştırma yöntemi, yani aile çalışmalarıyla vaka kontrol çalışmaları arasındaki önemli farklardan bir tanesi de, aile çalışmaları ile penetransı yüksek ama toplumda nadir görülen genetik varyasyonları bulma şansınızın yüksek olmasıdır. Vaka

kontrol ilişki çalışmalarında ise, penetransı düşük de olsa, toplumda sık görülen polimorfizmler tanımlanabilmektedir. Örneğin söz konusu genetik varyasyonun toplumda görülme sıklığı % 10–20 ise, bu sıklığın %40'a çıkmasının getirdiği riskten bahsedersiniz. Bu tür polimorfizmlerin penetransı düşük, ama toplamda görülme sıklığı göz önüne alındığı zaman, toplum için getirdiği risk çok daha yüksektir.

Bir pozitif ilişki tanımlandığı zaman, bunu açıklayabilecek nedenlerin tümü gözden geçirilmelidir: saptanan gerçek bir ilişki olabilir ya da tamamen tesadüfen oluşmuş veya sistematik bir nedene bağlı hatalardan kaynaklanabilir. Saptanan ilişki gerçek fonksiyonel ilişkiyle bağlantı dengesizliğinde olabilir. Bu nedenle bir ilişki tanımlandığında, bunun biyolojik anlamının da gösterilmesi gereklidir ve elde edilen sonuçların fonksiyonel analizlerle de desteklenmesi yararlı olacaktır. Tanımlanan polimorfizmin klinik uygulamalardaki yerinin belirlenmesi ise daha uzun zaman gerektirmektedir.

Bir Katılımcı: Ailevi Akdeniz ateşi olan hastalardaki tecrübelerimize göre, soyağaçlarını çıkartıyorken hastalar çoğunlukla en fazla birkaç nesil ötesini hatırlayabiliyorlar. En fazla dedelerine kadar uzağa gidebiliyorlar akrabalarını tanımlamakta. Çok uzak akraba dedikleri, mesela ikinci derece bir akrabalık olabiliyor. Sözünü ettiğiniz Behçet'li ailenin aile öyküsünü o kadar kuşak geriye giderek nasıl çıkartabildiniz?

Ahmet Gül: Birkaç nesile birden ulaşabilmek, genellikle uzak akrabalarda hastalık tanımlama sayesinde oluyor. Eğer ikinci, üçüncü dereceden kuzenlerde hastalık varsa, soyağacını çizdiğinizde karşınıza üç-dört nesil bilgisi çıkabilmekte. Bu nedenle uzak akrabalarda hastalık ayrı önem taşımakta. Bununla beraber, ailelerden bilgi alırken hatırlamadaki hatalar önemli bir yanılgı nedeni olabiliyor. Biz bilinen hastalardan yola çıkıyoruz. Eğer konmuş tanı varsa çok kıymetli, ama bazen tanı konmamış, fakat bütün belirtileri tarif eden bireyler olabiliyor. Aile çalışmalarında sıkıntılardan bir tanesi, tanı koymak için kullandığımız kriterlerin aile içerisinde kullanımının zorluklar göstermesi. Örneğin bir Behçet ailesinde öyle bireyler var ki, oral aftı var, paterji testi pozitif ve artriti var. Uluslararası çalışma kriterlerine göre hasta Behçet değil, ama biliyoruz ki bu aile bireyi Behçet hastası. Dolayısıyla aile içinde eksik belirtili ve kriterleri karşılamayan bireyleri bulma şansınız fazla. Ama ciddi bir analiz yapılacaksa, geriye dönük, hastayı hasta olarak tanımlama, yani bireyi *affected* olarak niteleme bir kesinlik içermek zorunda. Bu kesinliğe ulaşamıyorsak bunları bilinmeyen (*unknown*) kabul edip analizi ona göre yapıyoruz. Bu sayede onlardaki genotip bilgisini de analizlerde değerlendiriyoruz, ama onun hasta veya sağlıklı olma durumunu değerlendirmiyoruz.

KAYNAKLAR:

1. Guttmacher AE, Collins FS. Genomic medicine-a primer. N Engl J Med 2002; 347: 1512-20.
2. Cardon LR, Bell JI. Association study designs for complex diseases. Nature Rev Genet 2001; 2: 91-9.
3. Page GP, George V, Go RC, Page PZ, Allison DB. "Are we there yet?": Deciding when one has demonstrated specific genetic causation in complex diseases and quantitative traits. Am J Hum Genet 2003; 73: 711-9.
4. Knight JC. Regulatory polymorphisms underlying complex disease traits. J Mol Med 2005; 83: 97-109.
5. Clark AG. Finding genes underlying risk of complex disease by linkage disequilibrium mapping. Curr Opin Genet Dev 2003, 13:296–302.
6. Carlson CS, Eberle MA, Kruglyak L, Nickerson DA. Mapping complex disease loci in whole-genome association studies. Nature 2004; 429: 446-52.
7. Zondervan KT, Cardon LR. The complex interplay among factors that influence allelic association. Nat Rev Genet 2004; 5: 89-100.
8. Ewens WJ, Spielman RS. What is the significance of a significant TDT? Hum Hered 2005; 60: 206-10.
9. Evangelou E, Trikalinos TA, Salanti G, Ioannidis JP. Family-based versus unrelated case-control designs for genetic associations. PLoS Genet 2006; 2: e123.
10. Freedman ML, Reich D, Penney KL, et al. Assessing the impact of population stratification on genetic association studies. Nat Genet 2004; 36: 388-93.
11. Gehring NH, Frede U, Neu-Yilik G, Hundsdoerfer P, Vetter B, Hentze MW, Kulozik AE. Increased efficiency of mRNA 3' end formation: a new genetic mechanism contributing to hereditary thrombophilia. Nat Genet 2001; 28: 389-92.
12. Knight JC, Keating BJ, Rockett KA, Kwiatkowski DP. In vivo characterization of regulatory polymorphisms by allele-specific quantification of RNA polymerase loading. Nat Genet 2003; 33: 469-75.
13. Rebbeck TR, Spitz M, Wu X. Assessing the function of genetic variants in candidate gene association studies. Nat Rev Genet 2004; 5: 589-97.

GENETİK İLİŞKİ ÇALIŞMALARINDA KARŞILAŞILAN SORUNLAR

M. Tevfik Dorak

Konuşmamı Dr. Gül'ün konuşmasının devamı gibi görebilirsiniz. Genetik ile hastalık ilişkileri konusu üzerinde devam edeceğiz ancak bu sefer daha çok bizleri bekleyen sorunlar hakkında konuşacağız. Konuşma başlıklarımızın uzun bir listesi mevcut ve de şunları kapsıyor:

- Çalışmanın toplum bazında yapılıp yapılmaması, olgu ve kontrol örneklerinin aynı çalışma bazından seçilip seçilmemiş olması ve olgu örneklerinin *incident* veya *prevalent* olup olmaması (yeni tanı konulmuş olgular mı yoksa eski ve yeni tanı konulmuşların karışımı olup olmaması),
- *Population stratification* denilen, epidemiyolojide *confounding by ethnicity* diye de adlandırılan toplumun alt tabakalarının, aslın da olmayan bir ilişkiyi varmış gibi gösterecek komplikasyonlar yaratması,
- Türkçe'ye dengesiz bağlantı diye çevrilen *linkage disequilibrium* (LD) olayının gerçek bir ilişkiyi saklaması,
- İstatistik güç yetersizliği,
- Çoklu karşılaştırma sorunu,
- Bilinen diğer ilişkilerin dikkate alınmaması,
- Cinsiyete göre değişen bir ilişkinin gerektiği şekilde analiz yapılmadığı için gözden kaçması,
- Farklı genetik modellerin dikkate alınmaması,
- Bulguların değerlendirilmesinde aşırıya kaçılması
- Bulguların biyolojik temelinin olup olmaması,
- Yayınlamada tek yanlılık olayı (yalnız pozitif bulguların yayınlanması),
- Bulguların tutarlılığının tekrarlanan çalışmalarda gösterilmemesi.

Bu sorunların çoğu genelde epidemiyolojik çalışmalar için de geçerlidir ama içlerinde *population stratification* ve LD gibi sadece genetik çalışmalara özgü olanlar da mevcuttur. Çoklu karşılaştırmalar ve buna bağlı olarak şans sonucu gerçek dışı ilişki bulma sorunu da özellikle HLA gibi çok allelli lokuslar için dikkatli olunması gereken bir olaydır. Klasik epidemiyolojide bir

ilişki araştırılırken, o hastalıkta bilinen bütün diğer ilişkiler dikkate alınmaya çalışılır. Bu her zaman mümkün olmayabilir ancak genetik çalışmalarda eğer elinizde DNA varsa hiçbir mazeretiniz olmayacağından, bilinen diğer ilişkilere de bakmak ve sonuçlarınızı topluca değerlendirmek zorunluluğu doğar. Bu yapıldığı takdirde, gen-gen etkileşimlerine bakma şansımız da olacaktır. Etki değişimi çok ihmal edilen bir konu. Eğer ilişki sadece bir cinste görülüyorsa (ki nadir bir durum değildir) her iki cinsi ayrı ayrı analiz etmeyip topluca grup olarak bakarsanız, var olan bir ilişkiyi gözden kaçırmış olabilirsiniz. Bu konuşmada ağırlıklı iki temel sorun ele alınacak. Bunlardan ilki *population stratification* olacaktır. Türkiye gibi nisbeten karışık toplumları ilgilendirmektedir. Diğer ana sorun olan LD ise özellikle HLA kompleksi gibi genomun çok kalabalık bölgelerini özellikle ilgilendiriyor.

HLA kompleksinde 1 Mb başına ortalama 48 gen bulunmakta mevcuttur. Bu, genomda en kalabalık, yani en yüksek gen yoğunluğu olan bölge demektir. Böyle bir bölgede tek bir gene bakıp (örneğin HLA) bir ilişki bulununca, sanki o gen direkt olarak hastalıkla ilişkiliymiş gibi bir sonuca varmak mümkün değildir. Ancak bu, her gün de gördüğümüz bir durumdur. Bu durum için her zaman lösemide bulunan HLA-DP ilişkisini örnek gösteririm. Bu ilişki, HLA-DP ile olduğu için hastalıkta immün sistemin ve dolayısıyla infeksiyonların rolü olduğu şeklinde değerlendirilip yayınlanmış bulunuyor. Bu bulguyu ve varılan sonucu ciddiye almak mümkün değildir, çünkü HLA-DP'nin çok yakınında, çok kuvvetli LD içinde bulunduğu ve lösemi ile doğrudan ilişkili başka genler mevcuttur (örneğin retinoik asit reseptörü). Çok yakındaki diğer aday genlere bakmadan, HLA-DP ile bir ilişki bulundu diye hemen bulunan ilişkiyi HLA-DP'ye bağlamak uygun değildir. Unutmayalım ki, konjenital adrenal hiperplazi ile de önce bir HLA-B47 ilişkisi tanımlanmıştı, ancak sonunda hastalığa yol açan 21-hidroksilaz geninin HLA kompleksi içinde olduğu bulundu. Olay tamamen LD ile ilgiliydi, çünkü 21-hidroksilaz geninin hastalığa neden olan mutasyonu HLA-B47 haplotipinde bulunup dolayısıyla B47 ile LD gösteriyordu.

Population stratification toplumda farklı tabakalar olmasını tanımlıyor. Bu tabakalar farklı allel frekanslarına ve farklı hastalık insidanslarına sahip ise, bu karışımdan dolayı olmayan bir ilişki varmış gibi ortaya çıkabilir. Olay tamamen matematiksel bir komplikasyondur. Bunun klasik örneği, Amerika'da yapılmış immünoglobülin Gm alttipleri ile diyabet ilişkisini araştıran bir çalışmadır. Toplumda Kızılderililer ve beyaz Amerikalılar'ın olması ve bu iki grubun farklı allel ve hastalık frekanslarına sahip olması (aslında olmayan bir ilişkinin bulunmasına) neden olmuştur. Bunun nedeni olgu grubunda hastalık insidansı daha yüksek olan tabakanın, kontrol grubunda ise diğer tabakanın daha fazla temsil edilmesi ve bu tabakalarda, allel frekansları da farklı olduğu için, yapay bir farklılığın ortaya çıkmış olmasıdır. Bu sorunun önlenmesi icin, olgu ve kontrol seçimi sırasında örneklerin anne-baba ve hatta daha önceki nesilin sorgulaması yapılarak etnik kökenin tam olarak belirlenmesi denenebilir. Deneysel düzeyde ise bakacağınız aday hastalık genleriyle ilişkili olmayan, yaklaşık 30 adet birbirinden bağımsız genetik polimorfizm tiplendirmesi ile, olgu ve kontrol gruplarının gerçekten birbirinden farklı olup olmadığına bakılabilir. Bu dataya sahipseniz, özel yazılımlar kullanarak, toplumda tabakalar varsa bile bu durum dikkate alınarak, istatistik analiz yapmak mümkündür. Türkiye'de özellikle İstanbul'da yapılacak genetik çalışmalarda bu tür bir yaklaşım çalışmanın değerini ve sonuçların geçerliliğini mutlaka artıracaktır.

Son olarak genetik modeller üzerinde durmak istiyorum. Genetik çalışmalar gerek klasik epidemiyolojik çalışmalar, gerekse daha eskiden yapılan serolojik düzeydeki HLA çalışmalarından farklıdır. Diğer çalışmalarda yalnızca bir faktörün varlığı veya yokluğu önemli iken (örneğin sigara içip içmemek), modern genetik çalışmalarda allellerin kaç kopya bulunduğu ve bir kopya ile iki kopya olması halinde, risk ilişkisinin nasıl bir değişiklik gösterdiğine de dikkat etmek gerekiyor. Serolojik tiplendirme döneminde yapılan HLA çalışmalarında, sadece pozitiflik dikkate alınıyordu. Yani heterozigotlarla homozigotlar eşit ağırlıkla bir araya konulup, tek bir grup olarak (pozitiflik oranı) değerlendiriliyordu. Bu şekilde yapılan analiz dominant modele uyar ve de ilgilendiğiniz hastalık için her zaman uygun değildir. Çoğu zaman zaten hangi modelin uygun olduğunu bile bilmeyebiliriz. O yüzden sadece dominant modele bakarsak belki de bulabileceğimiz bir ilişkiyi kaçırmış da olabiliriz.

Genetik modeller konusunda çok güzel bir yazı olan, 2002'de *Briefings in Bioinformatics*'te yayınlanmış CM Lewis tarafından kaleme alınmış derlemeyi okumanızı öneririm (internetteki sunumumdan direkt link ile ulaşabilirsiniz).

Burada o yazıdan alınmış bir tabloyu veriyorum. Göreceğiniz gibi basit bir 2x3 genotip frekans tablosunu hazırladığınız takdirde bütün genetik modelleri yani resesif, dominant, aditif ve multiplikatif modelleri incelemeniz mümkündür ve de incelememiz gerekir. Genetik ilişkilerde en fazla rastlayacağınız model aditif modeldir ve eğer bu model ile anlamlı bir ilişki bulduysaniz gen-dozaj etkisi de bulmuşsunuz demektir; bu da ilişkiyi daha anlamlı bir şekle sokar (Hill's kriterlerinden biyolojik gradyan etkisine eşdeğer olması nedeniyle). HLA çalışmalarında çok sık yapılan bir başka hata da istatistik gücü artırdığı ve daha küçük bir p değeri verdiği için kromozom başına hesaplanan allel frekanslarının karşılaştırılmasıdır. Tablodan göreceğiniz gibi, bu multiplikatif modele karşılık gelir, diğer bir deyişle heterozigotlardaki *odds ratio*, homozigotluğa geçince katlanarak artıyorsa geçerlidir. Eğer böyle bir durum yoksa bu karşılaştırmayı yapmamalısınız. Yapacaksanız da örneğinizin Hardy-Weinberg dengesi gösterdiğinden emin olmalısınız. Çünkü artık genotipleri değil, onları parçalayarak allelleri tek tek kullanıyorsunuz. Allelleri bu şekilde tek tek kullanabilmeniz icin önce onların birbirinden bağımsız olduğunu; genotipe bağımlı olmadığını göstermeniz gerekir. Bu da Hardy-Weinberg dengesine bakılarak yapılır. Örneğin belli bir homozigot genotip toplumda beklenenden daha az sayıda bulunuyorsa genotipler arasında bağımsız bir ilişki yok demektir, o zaman da multiplikatif model testi olarak allel frekansı karşılaştırması yapamazsınız. Bunu yaptığınız zaman bulacağınız *odds ratio* hakkında da dikkatli olmanız gerekir. Artık bir allelin varlığı halinde o kişinin riskinin ne kadar arttığı söz konusu değildir, çünkü kullandığınız birim kişi değil kromozomdur. Yine de *odds ratio* ilişkinin kuvveti hakkında bir fikir verecektir.

Bu konuşmada kullandığım sunumu internette bulabileceğinizi hatırlatıyorum (http://www.dorak.info/epi).

(a) Genel genetik modeli için tam genotip tablosu

	AA	AB	BB
Vakalar	a	b	c
Kontroller	d	e	f

(b) Dominant model: allel B riski artırıyor

	AA	AB+BB
Vakalar	a	b+c
Kontroller	d	c+f

(c) Resesif model: riskin artması için allel B'nin iki kopyası gerekli

	AA+AB	BB
Vakalar	a+b	c
Kontroller	b+e	f

(d) Multiplikatif model: AB riskinde r kat artış, BB riskinde r^2 artış. Genotipe göre değil, allellere göre analiz edilmiş

	A	B
Vakalar	2a+b	b+2c
Kontroller	2d+e	e+2f

(e) Aditif model: AB riskinde 4 kat artış, BB riskinde r^2 artış. Genotipler Armitage eğilim testiyle analiz edilmiş

	AA	AB	BB
Vakalar	a	b	c
Kontroller	d	e	f

KATKILAR

Bir Katılımcı: Dr. Gül'ün modülüne dönersek, oradaki o paternler açısından çok bir şey değişecek mi? Yani FMF'te var olan bir şeyler bir taraftan Behçet'te de varsa, kontrol grubunda ve hasta grubunda da aynı asosyasyonları mı bulacağız? Bu o zaman o hastalıklar için bu asosyasyonların önemini kaybettirecek mi?

Hasan Yazıcı: Önem kaybettirmeyecek, zaman kazandıracaktır.

Bir Katılımcı: Genetik hastalıklarla uğraşan ama genetikçi olmayan klinisyenlerin tartıştığı sorulardan biri kontroller, diğeri genetik analizde *inter-* ve *intra-observer* hataların ne kadar önemli olduğudur.

Hasan Yazıcı: Gerçekten çok önemlidir.

Tevfik Dorak: Bunlar daha çok *high-throughput* çalışmalarda gündeme gelen konular, ancak benim başımdan böyle bir durum geçmedi. Ben daha çok küçük tek merkezli çalışmalar yapıyorum ve böyle bir durumla pek karşılaşmıyorum. *SNP chip* çalışmalarında özellikle tam genomda %1 dahi genotipleme hatası yapsanız sonuçlarınız çok etkilenir. O yüzden çok sıkı kontrolleriniz olması gerekir. Mesela plaklarla çalışırken her plakta bir kaç tane aynı DNA'nın olması, aynı plaktan bir DNA'nın başka bir kuyucukta da bulunması gerekir. Bu şekilde ve diğer eşdeğerliklerin mutlaka kontrol edilmesi lazımdır. Negatiflerinin mutlaka negatif olması, genotipleme hatalarını % 1'in de altına düşürmeniz gerekir.

Bir Katılımcı: Bu stratifikasyon konusu bu kadar tartışıldığına göre şunu sormak istiyorum: Türkiye'de birçok bölgeden HLA verisi mevcut, Antalya'dan, Ankara'dan, hatta Van'a kadar. O yayınlardaki HLA dağılımı, gerçekten sorun yaratacak ölçüde bir farklılığı gösteriyor mu ya da oradaki dağılımın varyansını anlamak bu sorunu çözmemizi sağlar mı?

Tevfik Dorak: İstanbul'da yapılan çalışmalara bakarsanız, aynı şehirden iki kontrol örneğinin bile birbirinden anlamlı olarak farklı olabileceğini görürsünüz. Bu varyanstan haberimiz olması yararlı olur ve sorunu tanımamızı sağlar. Yoksa o varyansı kullanarak herhangi bir düzeltme yapmamız söz konusu değil.

Hasan Yazıcı: İlaç çalışmalarını iki bölümde ele alacağız. Etki ya da iyileştirme konusunu Dr. Hamuryudan, yan etki konusunu ise Dr. Ioannidis anlatacak. İlaç çalışmaları hakkında çok kısa bir şey söylemek istiyorum: İlaç çalışmaları esasında çok güzel bir enstrümandır ve 1950'de ortaya çıkmaları hiç de şansa bağlı değildir. O sırada tüm dünyada bilimsellik değer kazanıyor ve kontrollü ilaç çalışması gibi bir enstrüman ortaya çıkıyor. Herhangi bir enstrüman bu; ölçmek için olan bir şey ama son yıllarda bu enstrümanla ilgili çok büyük sorunlar olduğu kanısındayım ben.

İLAÇ ÇALIŞMALARINDA ETKİ

Vedat Hamuryudan

Profesör Doktor, İstanbul Üniversitesi, Cerrahpaşa Tıp Fakültesi, Romatoloji Bilim Dalı

Randomize kontrollü çalışmalar, ilaçların etkilerini görmek ve ruhsat almak için altın standarttır. Önce etki (*efficacy*) ve etkinlik (*effectiveness*) tanımlarını açıklamak istiyorum. Kontrollü çalışmalar etki yani *efficacy*'yi, yani ilacın o kontrollü çalışmanın yapılmış olduğu gruptaki etkisini gösteriyor. Etkinlik (*effectiveness*) ise daha uzun dönemli, sadece seçilmiş bir grup ile değil, tüm hasta grubunda uzun gözlemleri gerektiriyor. Randomize klinik çalışmalar nisbeten kısa süreli ve pahalı çalışmalardır. Kesin kuralları vardır. Örneğin, her hastayı almayıp, sonuçlar homojen olsun diye bazılarını dışlayabilirsiniz. Günlük hayattaki gibi, ilaç dozları daha önceden bellidir ve çok rahat doz ayarlaması yapılamaz. Seçilen etki değerlendirme ölçütleri her zaman gerçek hayatta karşılaştığımız sorunlarla ilgili olmayabiliyor. Bulunan anlamlı bir sonuç, her zaman günlük klinik pratik için anlamlı olacağı anlamına gelmiyor.

Romatolojide randomize kontrollü bir çalışma yapmak o kadar kolay değildir. Bir kere, hastalıklar nonhomojendir. Örneğin, romatoid artrit seropozitif, seronegatif, erozif, nonerozif olabilir. Hastalığın gidişi çoğu zaman tam kestirilemez; örneğin, Behçet üveiti atak ve remisyonlarla seyreder ve iyi bir değerlendirme için uzun süre takip gerektirir. Romatoid artritli hastaları izlediğiniz zaman 60 ayda metotreksat, diğer DMARD'lardan üstün çıkıyor. Ancak bu gözlemsel çalışmayı 1 yılda kesecek olursanız, altın, sülfasazalizin, d-penisillamin ile eşit çıkıyor. Diğer bir deyişle, kısa gözlem süresinin böyle bir sakıncası mevcuttur. TNF inhibitörleri ile romatoid artritte yapılan kontrollü çalışmaların hasta alım kriterlerine baktığımızda, Amerika'da bir üçüncü basamak merkezde takip edilen 140 hastanın sadece % 5'inin bu kriterleri doldurduğunu görüyoruz. Başka bir örnek olarak bifosfonatlar ile ilgili çalışmalar verilebilir. Risedronatın kalça kırığını önlediğini gösteren bir çalışmada çalışmaya girme kriterlerinden biri T skorunun -4 olması, diğer bir başkası ise bir risk faktörünün bulunmasıdır. Bu çalışmadan risedronatın kalça kırığını azalttığı sonucu çıkmış. Bu çalışmanın sonuçları sadece, ileri osteoporozu olan hasta grubu için geçerli iken bizler, klinikte bütün hasta gruplarına uyguluyoruz. Bir çalışmada 4 büyük romatoloji dergisinde çıkan reklamların % 95'inde hata bulunmuştur. En önemli hata çalışma sonuçlarının çalışma yapılan hasta gruplarından başka gruplara uyarlanması olmuştur.

Bir yayında lupusta yapılan bütün randomize kontrollü grup çalışmaları toplanmıştır. Konuyla ilgili 94 tane çalışma ele alınmıştır. Doksandört lupus çalışmasının 37 tanesinde renal tutulum mevcuttur. Çift kör, tek kör, çapraz (*cross-over*) olan, plasebo kontrollü yapılan çalışmalar yer almıştır. Lupusta renal tutulum olmayınca biraz daha fazla plasebo kullanabilir, ancak nefrit grubunda plasebo kullanımı hemen aşağıya iniyor. Değerlendirme ölçütleriniz 5 tane olduğu zaman daha çalışmaya başlarken bir tane olumlu sonuç bulmanız % 20 olarak hesaplanıyor. Romatolojide en çok yapılan klinik çalışmalar, plasebo kontrollü çalışmalardır. Ya mevcut bir ilaca plasebo eklenir ya da mevcut aktif ilaç kesilir. Klinik çalışmada hastanın iyiliği, menfaati toplum menfaatinin üstünde olmalıdır. Bu ana ilkeye ek olarak, çalışmaya giren hastaların en iyi tedaviyi almalarını garanti etmeliyiz. Bir hastalık için etkili bir tedavi varsa, bu tedavi kısmi bile olsa, o zaman burada plasebo kullanamazsınız. Hastayı belli bir süre plaseboya maruz bırakmanız geriye dönüşümü olmayan bir takım değişiklere yol açacaksa, yine plasebo kullanamazsınız. Romatoid artriti ele alırsak, romatoid artrit için metotreksat iyi bir tedavi midir? Öyleyse plasebo kullanamazsınız. Üç ay plasebo kullanıp ondan sonra tekrar aktif ilaca dönmek isteseniz de bunu yapamazsınız. Peki, plaseboyu nasıl kullanabiliriz? Etanerseptle JRA'da yapılmış bir çalışma mevcuttur. Önce bütün hastalara etanersept verilmiş, etki görülmüş, ardından hastaları iki gruba ayırmışlar. Bir grup etanersept enjeksiyonu ile, diğer grup plasebo enjeksiyonu ile tedaviye devam etmiş. Plasebo enjeksiyonu alanlarda aktivasyon olunca açık olarak etanersept vermişler. Başka bir çalışmada bir gruba metotreksata ek olarak plasebo infüzyonları yapılmış, diğer gruplara da metotreksata ek olarak farklı dozlarda infliksimab infüzyonları yapılmış. Şu anda romatoid artrit çalışmalarında en çok kullanılan yöntem mevcut ilaca diğer denenen ilacın eklenmesidir.

Benim bahsettiğim çalışmalar hep *superiority* yani üstünlük hedefleyen çalışmalardır. Burada sıfır hipoteziniz A ilacının B ilacına eşit olmasıdır. Çalışmanın sonucunda A ilacını B ilacından daha üstün olduğunu bulduğunuz zaman bu bir üstünlük çalışması olur. Eşitlik (*equivalence*) çalışmalarında sıfır hipoteziyle alternatif hipotez yer değiştirmiş ve giderek daha aşağı olmama (*non-inferiority*) çalışmaları ortaya çıkmıştır. Bunlar niye çıkmıştır? Çünkü hep yeni bir takım ilaçlar geliştiriliyor, bu yeni ilaçlar aslında mevcut tedavide bir devrim yapmıyorlar, ufak bir takım faydalar sağlıyorlar. Tabii biyolojik ajanların romatolojiye getirdiklerini dışarıda tutuyorum. Ayrıca FDA gibi bir takım kuruluşlar klinik ölçüt olarak, ilacın gerçek hayatta neler yaptığını kanıtı ile görmek istiyor. Bu durumda bir üstünlük çalışması yapmak çok pahalı olur ve çok yüksek sayıda hasta gerekir. Eşitlik veya aşağı olmama için daha önceden etkisi gösterilmiş (yani kanıtlanmış) ilaç olmalıdır ve yeni geliştirilen ilaç da bir takım ek avantajlar getirmelidir. ACE inhibitörü olan enalapril, gayet güzel işe yaramaktadır. Onun yerine başka bir lapril getirmek istiyorsunuz. Buna *me too* ilaç diyorlar ve bir avantajı yok; çünkü zaten bir tane var. İkinci bir ACE inhibitörü getirmenin ne avantajı var? Ancak biri daha ucuz, daha seyrek alınıyor veya yan etkileri daha az ise, o zaman *me too* ilaç kavramından uzaklaşılıyor. Bir aşağı olmama (*non-inferiority*) sınırı tanımlıyorsunuz ve yorumunuzu şunlara göre yapıyorsunuz: Test ettiğiniz ilacın etkisi diğerine göre nerede çıktı? Güven aralığı nedir? *Non-inferiority* sınırının neresinde? Ancak bazen öyle bir sonuç çıkar ki geniş bir güven aralığı ile karşılaşırsınız. Bu *non-inferiority* çalışmalarının da kendine göre bir takım problemleri mevcuttur.

Öncelikle deney duyarlılığını tanımlayalım. Bir üstünlük çalışması yaptınız, güzel yürüttünüz ve çalışma sonucunda test ilacının öteki ilaçtan daha iyi olduğunu gösterdiniz. Bu deney duyarlılığı (*assay sensitivity*) iyi olmuş bir deney anlamına gelir. Ancak önceki oturumlarda Tip I-Tip II hatalardan söz edildi. Üstün olmasına rağmen ilacın üstünlüğünü gösteremediniz. Bu sizin deney duyarsızlığınızdan kaynaklanıyor. Aşağı olmama (*non-inferiority*) veya eşitlik (*equivalence*) çalışmasında ise siz eşitliği gösterdiniz. Ancak bu, kötü yapılmış ve kötü yapılması dolayısıyla gerçekleşmemiş bir üstünlük çalışması kadar değer taşır, değeri daha fazla değildir. Tedaviye kalkışma (*intention-to-treat*) analizi yapıyorsunuz. Bu analiz ile esasında etkisiz bir ilacı daha etkili bir hale getiriyorsunuz. Peki, "sadece çalışmayı bitiren hastalarla analiz yapalım" dediğiniz zaman sonuç her iki tarafa doğru da kayabilir. Bir *non-inferiority* veya *equivalence* çalışması yapılıyorsa, her iki analizin de kullanılması önerilir. Bir *non-inferiority* veya *equivalence* çalış-

masında hipotezden dolayı çalışmaya katılanlar, A ilacıyla B ilacının eşit olduğunu bildiklerinde ne kadar kör olabilirler? Önemli sorunlardan bir tanesi de bu *non-inferiority* sınırını göstermektir. Örneğin, 2005 Kasım ayında, lupus nefritli hastalarda yapılmış açık bir *non-inferiority* çalışması yayımlandı. Mofetil mikofenolat ve siklofosfamid karşılaştırılmış. Sonunda *intention-to-treat* analizde mikofenolat mofetil siklofosfamidden daha üstün çıkmıştır. Daha sonra bu makaleye gelen bir mektup yayımlandı. Mektubun yazarı çalışmada siklofosfamid ile alındığı söylenen cevap oranının % 6, halbuki bu güne kadar yapılan çalışmalara göre beklenen cevap oranının yaklaşık % 30 olması gerektiğini söylüyor. Mektubun yazarı, bunun nasıl açıklanacağını soruyor. Sonra da, siklofosfamidin etkisinin 5-6 yıldan sonra açıkça görüldüğünü ekliyor ve bunu 6 aylık bir çalışmayla nasıl gösterdiklerini soruyor. Yani bir klinik çalışmayı okurken çok eleştirel yaklaşarak değerlendirmemiz gerekiyor.

Etkinin ölçülmesi: Etkiyi esas olarak iki şekilde değerlendiririz. Biri, klinik ana noktalardır. "Hasta nasıl, hastanın sağkalımı nasıl?" gibi soruların yanıtlarına bakarız. Etkinliğin ölçüldüğü olayın gerçekleşme sıklığına göre yüksek sayıda hasta ve uzun süre takip gerektirebilir. Bunlar tek başına incelenebildiği gibi (örneğin ölüm, son dönem böbrek yetersizliği veya majör kanama) tek tek oldukları zaman sayı azsa hepsi bir arada (kompozit) olarak incelenebilir. Örneğin COX2 inhibitörü çalışmalarında peptik ülserin üç komplikasyonunun (perforasyon, obstrüksiyon, kanama) birden ölçülmesi gibi. Ancak bazen bu sonuçlardan klinik anlam çıkarmak zordur. Örneğin bir çalışmada aspirin kullanımının miyokard infarktüsü, inme ve kardiyovasküler ölümden oluşan kompozit nedenleri % 18 azalttığı gösterilmiştir. Ancak tek tek analiz ettiğimiz zaman aspirinin kalp krizi riskini % 40 azaltırken, kardiovasküler ölümlere etkisiz olduğu, beyin kanaması geçirme riskini ise % 22 artırdığı görülmüştür. Yani kompozit değerlendirmeler klinikte her zaman geçerli sonuçlar sağlamaz. İkinci etki değerlendirme ise *surrogate* denilen ve temsil eden olarak Türkçe'ye çevirebileceğimiz şekildir. *Surrogate* olarak nitelendirilen son noktaların, klinikte geçerli son noktaların bir yansıması olması gerekir. Örneğin, hipertansiyon, kardiyovasküler mortalite, miyokard infarktüsü ve inme için geçerli bir *surrogate* olmalıdır. Tansiyonu kontrol ediyorsanız ölümü, kardiyovasküler olayları v.b. azaltırsınız. Ancak *surrogate* ölçütlerden elde edilen sonuçları klinisyenler çoğunlukla abartır. Burada endüstrinin de bir katkısı olabilir. Örneğin, kırık riskinin gösterilmesinde kemik yoğunluğu bir *surrogate* gösterge olarak kullanılır. Hepimiz biliyoruz ki florür kemik yoğunluğunu yükseltiyor ama kırık riskini de artırır. Romatoid artritte bugün en çok seçilen marker radyolojik erozyondur. Anti-TNF ilaçlar erozyon gelişimini azaltır. Buradan erozyonun önlenmesinin uzun dönemde deformite ve sakatlığı önleyeceği sonucunu çıkartıyoruz. Ancak bunun yanıtı henüz bulunmamıştır.

Etki göstermede önemli bir konu subgrup analizidir. Klinik çalışmada hasta grubundan seçilmiş ortak özelliği olan bir alt grubun sonuçlarını irdelemek çok önemlidir, ancak çalışmanın daha başında bu analizin hangi alt grupta yapılacağının tanımlanması gereklidir. Bu gruplar az sayıda olmalı ve sonuçla ilişkili olarak tanımlanmalıdır. Pratikte bir çok subgrup analizi yapılır ve bu analizlerden sadece işe yarayanı bildirilir. Gerçekte yapılan tüm diğer subgrup analizlerinin de bildirilmesi gerekir. Bir de subgrup *rescue* vardır: Bir çalışma yapılmış, ama o çalışmadan etkili bir sonuç çıkmamış. Sonra bir subgrup analizi ile etki çıkmış. 2006'da yapılan glukozamin çalışmasının sonuçlarına göre, glukozamin genel olarak, yani tüm grupta etkisiz bulunmuş, ancak bir subgrup analizi yapılacak olsa orta ve şiddetli ağrısı olan hastalarda etkili olabildiği bildirilmiştir. *Lancet*'te 2001'de glukozaminle ilgili çıkan bir yazıda, 3 yıl boyunca ilaç kullanıldığında, plasebo kullananlarda eklem mesafesi azalmasının 0,31 mm, glukozamin kullanımında 0,06 mm olduğu belirtilmiştir. İstatistik olarak bu anlamlı bir sonuçtur. Peki hastaların kliniğine baktığımızda ne ifade ediyor? Bu hasta 65 yaşındaysa ona, "10 yıl kullanırsan senin eklemin 0,75 mm daha az çökecek," mi diyeceksiniz? Bu sonucun klinikte bir önemi var mı? Ağrısı azalacak mı? Bilinmiyor, ama sanki böyleymiş gibi kabul görüyor.

2001'de yapılan bir ACR toplantısında tümü endüstri tarafından desteklenen 45 randomize kontrollü çalışma anlatıldı; çalışmaların 45'te 45'i etki bildirmişti. Asıl konumuz yaptığımız klinik çalışmayı okuyucuya nasıl bildireceğimizdir. Consort Bildirgesi'nden bahsetmek istiyorum.

Bir klinik çalışmayı bir tıp dergisine vereceğiniz zaman Consort kurallarına uymanız gerekir. Uluslararası saygın dergiler yayınlanacak makalelerin bu kurallara göre hazırlanmasını önkoşul kabul ederler. Ayrıca 1 Temmuz 2005'ten sonra başlanmış olan bütün klinik çalışmaların yayınlanabilmeleri için bir veritabanına kaydolması gerekir. Çalışmanızı veritabanına kaydedip bir numara alıyorsunuz ve o numara olmadan dergiler çalışmanızı yayınlamıyorlar. Bu arada bilim adamları da veritabanına girip, kendi alanlarında kaç tane çalışma yapıldığını görebiliyorlar. Ek olarak, bir takım *free access* internet dergileri ortaya çıktı. Klinik çalışma etkisiz bulundu ise ve başka bir dergi okuyucunun ilgisini çekmez diye yayınlamıyorsa, bu dergilerde yine hakem değerlendirmesinden geçip, etkisiz de olsa yayınlanıyor. Yani gelecekte bütün klinik çalışmalar negatif sonuçlansalar bile yayınlanabilecekler. Okuyucunun onlara da ulaşabilmesini çok yararlı buluyorum.

Her şey etki değildir. Bir sonraki konuşmada yan etkii anlatılacaktır. Bazen etki, yan etkide gizli olabiliyor. Örneğin, elinizdeki bir ilacın koroner kalp hastalığına gayet iyi geleceğini öngörüyorsunuz. Çalışmayı yapıyorsunuz ancak maalesef sizi hayal kırıklığına uğratıyor. Hastalarda pek fazla etki göstermiyor ama birkaç yan etkiden bahsediliyor; bunlardan biri de ereksiyonda artış olması. Hepimiz biliyoruz, bu yan etkiden erektil disfonksiyon tedavisinde devrim yaratan bir ilaç çıkıyor.

KATKILAR

Hasan Yazıcı: Subgrup analizini yapmak genellikle günahlardan bir tanesidir ancak hangi yönde yapıldığı önemlidir. Randomize kontrollü çalışmanın amacı varsayımı çürütmektir. Her ilaç iyi değildir ve biz bunu kanıtlamaya çalışıyoruz; bilimsel yöntem budur. Bunu kanıtlayamazsak o zaman ilacın iyi olduğu kanıtlanıyor. Doğal olarak, subgrup analizini kendi ilacının iyiliğini kanıtlamak yönünde yaptığın zaman bu bilimsel yöntem olarak kötüdür; ama subgrup analizini kendini çürütmek yönünde yaparsan o zaman tam tersine sevaptır. Bir örneğini vereyim: bir NSAID ile çalışma yapıyoruz. Bir NSAID'nin öbür NSAID'ye üstünlüğü gözüküyor. Ağır hastalar ve hafif hastalar gruplandırıldığında ağır hastalar arasında hiçbir fark görülmüyor. Üstünlüğün sadece hafif hastalar arasında olduğu ortaya çıkıyor. Bunu görünce çalışmayı destekleyen itiraz ediyor ve subgrup analizlerini reddediyor. Sonuç olarak, subgrub analizinin yapılma nedeni önemlidir.

Yaşar Karaaslan: Randomize kontrollü çalışmaların kısıtlılıklarını Prof. Hamuryudan çok güzel anlattı. Bu çalışmalar çok az sayıda hasta içerir, pahalıdır ve kısa sürelidir. Klinikte ilaçları kullandığımız zaman hepimizin bir gözlemi oluşuyor ve biz, randomize kontrollü klinik çalışmalarda söylenen etkiyi bulamıyoruz. Ancak bu verileri, yani kendi gözlemlerimizi bir şekilde değerlendirmek ve yazmak o kadar zor ki! Bunların toplanacağı bir yer olsa da gerçek hayattaki verilerimizi biraraya toplamamız mümkün olabilse. Herkes hastasını internete belli kurallar içinde yazar, belli şeylere bakar. Acaba dünyada böyle bir eğilim var mı? Yani tek tek klinisyenlerin gözlemlerinden bir şeyi sistematik olarak çıkarmak mümkün mü?

Vedat Hamuryudan: Biliyorsunuz bir sürü hasta veritabanı kuruldu; İsveç romatoid artrit veritabanı, Alman veritabanı gibi. Bu veritabanındaki hastaların belirli bir grubuna, diyelim A ilacı başlanıyor, A ilacını alanlar 1, 2 veya 3 sene izleniyor. Bu sürede kliniğe gelmiş, ilacı değiştirilmiş, başka bir ilaç eklenmiş olabilir. Komorbiditeleri, tüm verileri kaydediliyor. Buradaki verileri belirli bir süre sonra alıp, analiz ediyorsunuz. Bunlara uzun dönemli, gözleme dayalı çalışmalar deniyor. Bunlar çalışma hiyerarşisinde randomize kontrollü çalışmaların altında yer alıyor, ama esasında güzel yapılmış gözleme dayalı çalışmaların randomize kontrollü çalışmalar kadar değerli sonuçlar verdiğini söyleyen bir çok yayın var.

KAYNAKLAR:

1. Karassa FB, Tatsionni A, İoannidis JP. J Rheumatol 2003; 30: 979 – 984.
2. Pincus T, Sokka T. Ann Rheum Dis 2004; 63 (Suppl 2) ii32-ii39
3. Sokka T, Pincus T. Arthritis Rheum 2003; 48: 313 – 318.
4. Mcclung MR, Geusens P ve ark. N Engl J Med 2001; 344: 333 – 340.
5. van Winkelen P, van Denderen JS ve ark. Rheumatology 2006: 45: 1154 – 1157.
6. Lovell DJ, Giannini EH ve ark. N Engl J Med 2000; 342: 763 – 769.
7. Lipsky P, van der HeijdeDM ve ark. N Engl J Med 2000; 343: 1594 – 1602.
8. De Mets DL J Intern Med 2004; 555: 529 – 537.
9. Snapinn SM. Curr Control Trials Cardiovasc Med 2000; 1: 19 – 21.
10. Ginzler EM, Dooley MA ve ark. N Engl J Med 2005; 353: 2219 – 2228.
11. Karassa FB ve ark N Engl J Med 2006; 354: 764 – 765.
12. Schulz KF, Grimes DA. Lancet 2005; 359: 781 – 785.
13. Rothwell PM Lancet 2005; 365: 176 – 186.
14. Clegg DO, Reda DJ ve ark. N Engl J Med 2006; 354: 795 – 808.
15. Reginster JY, Deroisy R ve ark. Lancet 2001; 357: 251 – 256.
16. Fries JF, Krishnan E. Arthritis Res Ther 2004; 6: R505 – R513.
17. Altman D, Schulz KF ve ark. Ann Intern Med 2001; 134: 663 – 694.

İLAÇ ÇALIŞMALARINDA ZARAR

John P. A. Ioannidis

Programdaki konuşma için başta önerilen başlık "advers olayların bildirilmesi"ydi ancak gördüğünüz gibi başlığı zarar (*harm*) bildirimi olarak değiştirdim. Bence burada terminoloji önemlidir. Advers etki ya da advers olay terimleri teknik olarak doğrudur. Ancak çok teknik ve soyut terimler oldukları için bu kelimeleri kullanırken gerçek insanların başına gelen gerçek zararlardan bahsettiğimizi unutma eğiliminde olabiliriz. Consort Bildirisi'ne yapılacak ekler hakkında geçenlerde yaptığımız toplantıda panelistler olarak advers etki yerine zarar sözcüğünü uyarlamak ve kullanımını desteklemek gerektiğine karar verdik. Bu konuda kullanılan başka kelimeler de var. Örneğin güvenlik = *safety* yaygın kullanılan bir terim. Ancak güvenlik sözcüğünün yanlış izlenim uyandırma ihtimali de mevcuttur. Bu terimi kullandığımızda bir tedavinin hem etkili hem de güvenli olduğunu söyleyerek tedaviye iki kere prim vermiş oluyoruz. Böyle bir kullanım ciddi yanılgılara neden olabilir. Kullandığımız diğer terimlerden yan etki de sorunlu bir sözcüktür. Burada da ilaçlardan kaynaklanan zararların sanki çok önemli olmadığı izlenimi oluşuyor.

Randomize kontrollü çalışmaların 58 yıllık geçmişi içinde tıp literatürü zararları tam anlamıyla ihmal etmiş hatta yasaklı sözcük olarak görmüştür. Tipik düşünceye göre bir tedavi, aksi ortaya çıkana kadar genellikle güvenli kabul edilir. Bu bence yanlıştır. Tam tersi, bence bir tedaviyi aksini kanıtlayana kadar potansiyel zararlı olarak kabul etmek gerekir. Pek çok durumda da elimizde herhangi bir tedavinin zararsız olduğunu kanıtlayacak veri bulunmaz.

Tıp literatürü, zarar bildiriminin gözardı edilmesiyle doludur. Bir süre önce değişik dergilerde yayınlanmış bir grup randomize kontrollü çalışmayı değerlendirdiğimizde zararların bildirimi konusu oldukça fakirdi. Öyle ki, zararların bildirildiği toplam sayfa; yazarların çalıştıkları kurumu gösteren bölümün alanı kadar ya da ondan daha azdı. Yazılarda zararları bildirmeme eğilimi çok belirgindir. Yedi değişik bilim dalında çıkmış makaleleri tarayıp zararı,

a. klinik zarar,
b. laboratuara göre toksisite, ve
c. çalışmadan toksisite nedeniyle çıkış,

şeklinde 3 ayrı boyutta ne ölçüde bildirildiğini incelediğimizde çok az sayıda makalede herhangi bir boyutta % 50'nin üzerinde bildirim olduğu

gördük. Pek çok makale tüm boyutlarda düşük puan aldı. Doz karşılaştırması yapılmazsa, zararlar üzerindeki vurgu daha da sınırlı kalır. Örneğin, doz karşılaştırması yapıldığında araştırmacılar belli bir dozun etkinlik açısından aynı şekilde işe yarayabildiğini düşünürken, aynı zamanda zarar profili açısından da avantaj sağlayıp sağlamadığını görmek ister. Çalışmada değerlendirilen tedavi, önceden endikasyon almış bir tedavi ise, zarar bildirimi yine geri planda kalır ve yeterince ilgi görmez. Öte yandan eski bir ilacın farklı bir durum ve popülasyonda aynı toksisite profilini taşıyacağını garanti etmek mümkün değildir. Çalışmada etkinlik açısından anlamlı sonuçlar bulunduğunda, tedavinin verdiği zararlara yeterince dikkat edilmemiş olabilir. Son olarak yüksek *impact* faktörlü dergilerin zarar bildirimine daha fazla yer ayırdığına dair kanıt da yoktur.

Zarar bildiriminin ihmal edilmesiyle ilgili ampirik veri, ilaç çalışmalarının yanında ilaç dışı çalışmalardan da elde edilmiştir. Hatta onlarda durum çok daha kötüdür. Romatizmal hastalıkların ilaç dışı yollarla tedavisini konu alan çalışmalarda zarar bildirimini konu alan çok iyi bir çalışma vardır. Sonuçta zarar bildirimi açısından bu tür çallşmalarda da, hatta ilaç çalışmalarından daha da sorunlu bir tablo ortaya çıkmıştır.

İlaç dışı tedavileri konu alan diğer randomize kontrollü çalışmaların pek çoğunda da zarardan hiç söz edilmez. Mental sağlıkla ilgili tedavilerin değerlendirildiği randomize kontrollü çalışma makalelerine baktık. Mental sağlık alanındaki büyük bir veritabanını tarayarak, çok sayıda randomize kontrollü çalışmadan oluşan bir grup seçtik. Bu grupta davranış tedavisi gibi ilaç dışı tedavilerin değerlendirildiği çalışmalarda, ne gibi zararlar ortaya çıkabileceğine dair bir tek bilgi dahi verilmiyordu. Ortalama olarak, bir makale ilaç çalışması ise zararlar için 1/10 sayfa yer ayırıyor, ilaç dışı tedavi çalışması ise hiç yer ayırmıyor.

Burada sistematik derleme ya da meta analizlerde zarar sıklığı konusunda daha iyi bilgi edinilip edinilemeyeceği akla gelebilir. Durumun böyle olup olmadığına bakmak için Cochrane sistematik derleme veritabanını araştırdık ve en az 4000 hasta ile ilgili kanıt içeren 138 derleme bulduk. Sadece 25 tanesinde anlamlı, büyük ölçekli kanıt olarak kabul edilebilecek advers olay verisi vardı (en az 4000 hasta hakkında bilgi ile birlikte iyi tanımlanmış zarar verisi).

Zararlar konusunda büyük ölçekli veri eksikliği bilgisi değişik şekiller aldı. Örneğin en az 4000 hasta hakkında bir kısım kantitatif veri bildirilmiş olabilir ancak advers olaylar nonspesifik olarak bildirilir (değişik advers olayların birleşik sayısı gibi). Ne tür advers olayların kastedildiğini bilmiyorsanız verilen bilgi çok işe yaramaz. Bir başka örnek zararın şiddet ve tanım belirtilmeden bildirilmesidir. "Şu sayıda insan diyare oldu," dediğimizde, birkaç yumuşak dışkılamayı mı kastediyoruz, 5 aydır diyare mi olduğu, yaşam kalitesinin çok mu bozulduğu, hastaneye yatması mı gerektiği anlaşılmaz. Son olarak, derlemelerin bir kısmında 4000'den fazla hasta hakkında etkinlik verisi bildirilirken spesifik ya da nonspesifik zarar verisi bildirilen hasta sayısı 4000'in altındadır.

Peki randomize kontrollü çalışmalardan elde edemediğimiz bilgiden kaynaklanan boşluğu randomize olmayan veriyle doldurabilir miyiz? Aslında bu güne kadar klinik pratikte yaptığımız budur. Şimdiye kadar olan görüş randomize kontrollü çalışmaların zararlar açısından çok aydınlatıcı olmadığıdır. Bir ilaç pazara çıkana kadar sadece ilacın faydasını ortadan kaldıracak derecede ciddi ya da çok sık görülen bir yan etki olmadığından emin olmaya çalıştık. Öte yandan doktorların gelecekte bu ilacın vereceği zararları pazarlama sonrası takip sırasında bildireceğini düşündük ve iyi planlanmış faz 4 çalışmalar olacağını öngördük. Ancak böyle çalışmaların da sayısı çok azdır. Sonuç olarak tedavilere bağlı zararların ne olduğunu öğrenmek için temelde gözlemsel çalışmalara güvenir hale geldik. Bazen gözlemsel çalışmalarda elde ettiğimiz verilerle randomize kontrollü çalışmalar birbirini tutar, bazen de tutmaz. Örneğin laparoskopik ve açık apendektomi sonrası yara infeksiyonu sıklığı konusunda hem randomize kontrollü çalışmalar, hem de gözlemsel çalışmalar vardır ve benzer sonuçlar vermişlerdir. Ancak aynı karşılaştırmayı damar ya da organ yaralanması konusunda yaparsak randomize kontrollü çalışmalarla gözlemsel çalışmalar arasında on kat gibi bir fark görülür. Yakın zamanda yaptığımız ampirik bir değerlendirmede, randomize kontrollü çalışmalarla gözlemsel çalışmalar arasında zarar bildiriminin farklı olup olmadığına baktığımızda orta düzeyde bir uyum gördük.

Peki, gözlemsel çalışmadaki verinin randomize kontrollü çalışmalardakine benzer olup ol-

madığını, randomize kontrollü çalışma yapılmış olsa aynı olup olmayacağını ya da tam tersini, yani elde gözlemsel veri olsa randomize kontrollü çalışmadakine benzer sonuçlar çıkıp çıkmayacağını nasıl anlayabiliriz? Bunu söylemek çok zordur. Pek çok durumda zararların ne sıklıkta ve ne şiddette olduğu konusunda bir sürü belirsizlikle karşı karşıya kalınacaktır.

Pek çok çalışma zararlar konusunda genel ve üstü kapalı laflar eder. İlacın "iyi tolere edildiğini" söyler. Bu kadar lafı söylemek için çalışma yapmaya gerek yok ki! Biz bu konuda daha fazla şey bilmek istiyoruz. "Çalışmadaki 22 hastada ağır ishal ortaya çıktı". Bu da oldukça sık kullanılan bir bildirim şeklidir. Peki ishal çalışmanın hangi kolunda oldu? Deney ilacında mı yoksa karşılaştırma kolunda mı? Bir başka yazıda "22 tedavi hastasında ağır advers olaylar görüldü" ifadesini okuyabiliriz. Burada bir hastaya dönüp, ağır bir yan etki görme ihtimalimiz % 0,1, % 1 ya da % 2'dir diyemezsiniz. Görülen yan etki nedir? Kanama mıdır, baş ağrısı mıdır, kaşıntı mıdır? Son olarak yukarıda anlatılmaya çalışıldığı gibi çalışmada "23 hastada ishal oldu" denebilir ancak bu sefer de ishalin şiddeti bildirilmemiş olur

Zararlar konusundaki bir başka hatalı yorum da, çalışmanın temel amacı olmamasına rağmen zararlarla birlikte bildirilen yalancı anlamlı istatistik sonuçlara odaklanmaktır. Çalışma sonucunda "tedavi kolunda diğer koldan daha sık görülen tek sorun diyareydi" yazılabilir ve tedavi kolunda $p<0,05$ gibi bir diğer, hatta tam p değeri dahi bildirilebilir (örn., $p = 0,002$).

Devamlı verilerle ilgili bilgilerin ekstrem değerler bildirilmeden sadece ortalama ve medyan gibi merkezi eğilim göstergeleriyle bildirilmesi de karışıklık yaratır. Oysa zararlar konusunda ekstrem değerler önemlidir. Örneğin, hemoglobin değerini alalım ve 100 hastadan birinin aplastik anemi nedeniyle hemoglobin değeri 3 olarak öldüğünü varsayalım. Yüz hastanın hemoglobin ortalamasına bu değeri koyduğumuzda ortalamanın 0,1 puan daha az çıkmasına neden olacak ve asıl sorun olan ekstrem olayı yakalamak mümkün olmayacaktır.

Zamanın yan etki konusunda belirleyici etken olduğu durumlarda yan etkinin ortaya çıkış zamanını anlamak önemlidir. Yan etkilerden bahsederken çalışmanın erken döneminde mi, geç mi yoksa orta vadede mi hatta çalışmanın takip süresinden sonra çıkıp çıkmadığını düşünmeliyiz. Tüm bunlar birbirinden tamamen farklı şekilde yorumlanabilir. Şunu kabul etmemiz lazım ki, eğer ana etkiye bakmaya yetecek kadar kuvvetli bir çalışma planlamamışsak zamana bağlı etkiler konusunda yapacağımız yorumlar daha zayıf olacaktır. Bazen zamana bağlı olayları bu şekilde incelemek mümkün olabilir. Bazen de daha önceki çalışmalar ya da başka kanıtlardan yola çıkarak zararların daha yüksek olasılıkla ya da daha şiddetli ortaya çıktığı dönemlere odaklanabiliriz; bu dönem çalışmanın ilk altı ayından sonra, ya da ilk altı ayı ya da tedavi başlangıcından sonraki ilk hafta olabilir. Böyle durumlarda tüm takip süresini ortalamaya vurmak, orada bulunabilecek zamana bağlı etkiyi seyreltecektir.

Başına bir tane olay gelen hastalarla birden fazla olay gelen hastaları ayırmamak da yanıltıcı yorumlara neden olabilir. Bazı katılımcılar aynı tür advers olayları tekrar tekrar, çalışma boyunca 3-5 kez yaşayabilir. Böyle bir durumda, olayın kaç kez tekrar ettiğini bilmek isteriz.

Son olarak, zararlar konusunda bildirilen verinin randomize edilen tüm hastalar üzerinden mi yoksa sadece bir alt grup üzerinden mi verildiğine dikkat etmek gerekir. Zararlardan söz edilirken tedavisi hedeflenenlerin (*intention-to-treat*) analizi yerine tedavi alanların analizini yapmak daha iyi olur diye düşünülebilir. Böyle düşünmekte bir hakikat payı da vardır çünkü bir hasta bir tedaviyi almamışsa, ortaya çıkan zararı o tedaviye bağlamak mümkün olmaz. Ancak istatistik açısından bakıldığında tedavisi hedeflenen hastaların analiz edilmesi daha avantajlıdır. Diğer taraftan tedavi alanlar diye bir alt grup tanımlamaya çalıştığımızda ortaya yeni seçenekler ve seçimler çıkacak, tanımlar manipüle edilebilir hale gelecektir. Örneğin tedavisiz kalmanın tanımı hiç almamış olmak mı, tek doz almak mı, tedaviyi 1 haftadan daha kısa süre kullanmak mı yoksa başka bir şey mi olacaktır?

İlaç dışı tedavi çalışmalarından kaynaklanan zararları yakalamak ve tanımlamak kendine özgü güçlükler taşır. Öncelikle ilaç dışı girişimlerin de zarar verebileceğini anlayarak işe başlamak gerekir. Pek çok cerrah hastalara zarar veremeyeceğine inanır. Psikoterapistler de böyle hisseder. Öte yandan görünürde etkili bir psikoterapi

seansı gören bir hasta seanstan sonra evine gidip intihar edebilir. İlaç dışı çalışmalarda, başta ortaya çıkabilecek zararların doğru tanımlanması olmak üzere çalışma tasarımı ve veri toplama yaklaşımı üzerine fikir üretmemiz gerekli. Bunun yanında durum, cerrah, öğrenme eğrisinin safhaları ve girişimin ne kadar standartlaştırılabileceğini de göz önünde bulundurmalıyız. İlaç dışı çalışmalar genellikle körleme olmadan yapılır ve bu durum ortaya çıkan zararların anlamlandırılmasında sorun yaratabilir. Ayrıca ilaç dışı girişim çalışmalarında zarar tanımı yapmanın zor olacağını da düşünmeliyiz, çünkü bu alışmalarda ilaçlarda olduğu gibi standart yan etki listeleri olmayabilir. Bu nedenle tanımlar dikkatli ve titiz yapılmalıdır. Nelerin yolunda gitmeyebileceğini baştan düşünmek gerekir.

Zamanlama da çok önemli olabilir. Cerrahi çalışmalarında sık görülen bir durum, araştırmacıların sadece ameliyat sırasında olanları bildirmesidir. Ameliyat başarılı ve sorunsuz geçmiş olabilir. Ancak belki hasta derlenme odasına döndükten uzun süre sonra ameliyatla ilgili ciddi bir komplikasyon geliştirmiştir. Ya da zarar kendini girişimden uzun süre sonra belli edebilir, bu durumda takibi ameliyattan sonraki birinci haftada sonlandırırsak çok önemli olabilecek geç komplikasyonlar konusunda hiç bilgi edinmemiş oluruz.

Sonuç olarak, bence hastalar üzerinde uyguladığımız girişimlerin ortaya çıkarabileceği zararlar konusunda daha fazla bilgi talep etmeliyiz. Geleneksel olarak bu bilgi azdır. Diğer bir deyişle, faydaları zararlarla dengelemeye çalışırken zararlar hakkında bildiklerimiz ihtiyacımız olanın yarısı ya da daha azıdır. Öte yandan hiç zarar vermeyen bir girişim muhtemelen iyileştirme açısından da pek etkili olmayacaktır. Hiç bir girişim tamamen güvenli değildir. Teşekkür ederim.

KAYNAKLAR:

1. Karassa FB, Afeltra A, Ambrozic A, Chang DM, De Keyser F, Doria A, Galeazzi M, Hirohata S, Hoffman IE, Inanc M, Massardo L, Mathieu A, Mok CC, Morozzi G, Sanna G, Spindler AJ, Tzioufas AG, Yoshio T, Ioannidis JP. Accuracy of anti-ribosomal P protein antibody testing for the diagnosis of neuropsychiatric systemic lupus erythematosus: an international meta-analysis. Arthritis Rheum 2006;54:312-24.

2. Lau J, Ioannidis JP, Schmid CH. Summing up evidence: one answer is not always enough. Lancet 1998;351:123-7.

3. Ioannidis JP, Gwinn M, Little J, Higgins JP, Bernstein JL, Boffetta P, Bondy M, Bray MS, Brenchley PE, Buffler PA, Casas JP, Chokkalingam A, Danesh J, Smith GD, Dolan S, Duncan R, Gruis NA, Hartge P, Hashibe M, Hunter DJ, Jarvelin MR, Malmer B, Maraganore DM, Newton-Bishop JA, O'Brien TR, Petersen G, Riboli E, Salanti G, Seminara D, Smeeth L, Taioli E, Timpson N, Uitterlinden AG, Vineis P, Wareham N, Winn DM, Zimmern R, Khoury MJ; Human Genome Epidemiology Network and the Network of Investigator Networks. A road map for efficient and reliable human genome epidemiology. Nat Genet 2006;38:3-5.

4. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schunemann HJ, Edejer TT, Varonen H, Vist GE, Williams JW Jr, Zaza S; GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ 2004;328:1490.

5. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. N Engl J Med 1992;327:248-54.

6. Ioannidis JP, Rosenberg PS, Goedert JJ, O'Brien TR; International Meta-analysis of HIV Host Genetics. Meta-analysis of individual participants' data in genetic epidemiology. Am J Epidemiol 2002;156:204-10.

7. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. BMJ 2006;333:597-600.

8. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. JAMA 2005;294:218-28.

9. Pan Z, Trikalinos TA, Kavvoura FK, Lau J, Ioannidis JP. Local literature bias in genetic epidemiology: an empirical evaluation of the Chinese literature. PLoS Med. 2005;2:e334.

İYİ KLİNİK UYGULAMALAR

Cem Özesen

Bu oturumda İyi Klinik Uygulamalar'ın ana hatlarının üzerinden geçiyor olacağız. İki kavramdan bahsedeceğiz: ilki *GCP* ya da *Good Clinical Practice* ya da *Good Clinical Research Practice*, diğeri ise bütün dünyada kabul edilen biçimi ile ICH GCP'dir. GCP, 1997'deki ilk ICH GCP'deki tanımına bakarsanız, bir etik ve bilimsel kalite standardıdır. Bu standart neyi belirler? İnsanların katıldığı çalışmalarda, çalışmaların tasarımı, yürütülmesi, kaydedilmesi ve bildirilmesinin standardını oluşturur. Bu standartlara uyulması sayesindeyse temelde iki yarar elde ediyoruz: Bunlardan ilki, bu çalışmalara katılan insanların haklarının, emniyetinin ve iyilik hallerinin korunması, ikincisi ise elde edilen verilerin güvenilir olmasıdır. Bu tanımı daha şematik görmeyi denediğimizde, GCP'nin bize çalışmaların tasarımı, yürütülmesi, kaydedilmesi, izlenmesi, denetlenmesi ve bildiriminin standartlarını sağladığını görürüz. Bunun sonucunda da, bu standartların uygulanması ile elde edilen veri ve bildirilen sonuçlar doğru ve inanılır kabul edilmektedir.

Eş zamanlı olarak ikincil hedefi ve sağladığı ikincil kazanım bu çalışmalarda yer alan insanların hakları, bütünlükleri, bilgilerinin gizliliği, sağlıkları ve iyilik hallerinin korunmasıdır; bunun güvencesini oluşturan GCP bazı ülkelerde yasal bir zorunluluk olmakla birlikte, bazı ülkelerde süreçler halen devam etmektedir. Bazı ülkelerde bir çalışma yapılacaksa GCP'ye uymak bir zorunluluk olarak geçmekte, kimi ülkelerde ise cezai müeyyideleri olan bir yasal çerçevenin içinde yer almaktadır.

Neden GCP'yi bilmek gerekir? GCP'yi bilmeden, GCP uygulamadan yapılan bir klinik çalışmanın sonuçları, bu sonuçların ulaşması beklenen noktalardan reddedilip geri dönebiliyor. Bilimsel çevreler GCP'ye uymadan yapılan klinik çalışmaların sonuçlarını reddedebiliyor ya da bir ilacın ruhsatlandırılmasına yönelik bir çalışma söz konusu ise, GCP uyumsuzluğu, ruhsatlandırma otoritelerinin o çalışmanın verilerini kabul etmemelerine yol açıyor. Bunun sonucunda, çalışmayı yapan araştırıcı da kişisel olarak etkileniyor ve artık inandırıcılığı, yaptığı işin güvenilirliği konusunda soru işaretleri ortaya çıkıyor. GCP'ye uymayan araştırıcı hem kendi inanılırlığını, hem de çalıştığı kurumun güvenirliliğini aşındırmış oluyor.

Doktor, Ares CRO, İstanbul

Yakın zamanlarda değil, bundan 7–8 yıl önce başıma gelen bir olayı örnek verebilirim; hala benzer şeylerin yaşanması mümkündür. Bir hekim arkadaşımız bir gün bana telefon ederek, makalesini bir dergiye göndererek yayımlanması için başvurduğunu ancak derginin çalışmanın etik kurul onayını sorduğunu ve alınmamış olduğunu öğrenince de makalenin reddedildiğini söyleyerek, etik kurul onayını şimdi nasıl alabileceklerini sordu. Çalışma bitmiş, makalesi yazılmıştı ve artık bunun etik kurul onayı olamazdı. Peki bunu sonuçları ne oldu? Bir dergi makaleyi reddettiyse başka bir dergiye göndermemiz mümkün değildir. Başka dergiye gönderse bile, verdiği zarar artık geri dönülebilir değildir. Peki ne zarar vermiş olabilir? Bir dergi Türkiye'den, bir üniversiteden gelen bir makalenin, etik kurul onaysız bir çalışmanın makalesi olabileceğini gördü. Bu önyargıyı besleyen bir şeydir; artık belki de o dergiye Türkiye'den giden bir başka makale de önyargıyla algılanacak. Sonuçta GCP'ye uymak hepimizin yapması gereken bir şeydir; yalnızca kişisel kazanım için değil, daha büyük düşünüldüğünde de işimize yarıyor. Klinik çalışmalar yalnızca ilaç çalışmaları demek değildir ancak işin bir ilaç çalışması boyutu da vardır. Klinik çalışmaların bir başka komponenti olan ilaç şirketleri için de bu konu son derece önemlidir. GCP'yi ilaç şirketleriyle koşut algılamak gerekir çünkü ilaç endüstrisi her zaman için mutlaka GCP'ye uyum istemektedir. İlaç endüstrisi yaptığı çalışmaların büyük bir bölümünü ruhsatlandırma amacıyla kullanmayı hedeflediğinden GCP uyumu söz konusu değilse ilaç şirketinin o çalışmalarını ruhsatlandırma verisi olarak kullanması mümkün değildir. Klinik çalışmalar çok pahalı girişimler olduğu için GCP uyumunun olmaması bir yandan dünya kadar paranın da sokağa atılması anlamına gelebilir. Öte yandan GCP'ye tam uyum deneklerin, insanların, çalışmaya katılanların hakları ve sağlıklarınının korunmasının yanı sıra, yaptığımız işin kalitesininin güvence altına alınmasını da sağlıyor. Sonuçta GCP'ye uyarak bir çalışma yapıyor olmak, yaptığınız çalışmanın bilimsel açıdan su geçirmez olması için de kullanmamız gereken önemli enstrümanlardan bir tanesidir.

GCP'nin ya da ICH GCP'nin gelişimine bakarsak, GCP'nin doğum yerinin Amerika Birleşik Devletleri olduğunun kabul edildiğini görürüz. GCP olarak kodladığımızda böyle olmakla birlikte, sürecin ABD'deki doğumunun öncesi de mevcuttur. Profesör Yazıcı'nın daha önce söylediği gibi, ilaç çalışmalarının geçmişi 150 yıla uzanıyor ancak bunun 2. Dünya Savaşı'na kadar olan bölümünde hiçbir kural yoktur. "Hiçbir kural yoktur," derken, gerçekten de hiçbir düzenleyici olmadan 2. Dünya Savaşı sonuna kadar geliniyor. Bunun bize kaybettirdiği çok şey var. Basit bir örnek verirsek, geçen yüzyılın sonlarında GCP kuralları işliyor olsaydı, biz parasetamolü (asetaminofen) kullanmak için 1958 yılına kadar beklemezdik. Parasetamolle yapılan ilk çalışmalarda, 19. yüzyılın sonlarında, bir makalede ilacın toksik olduğu belirtilince asetaminofenin ilaç olarak kullanımından vazgeçilmiştir. 1950'lerden sonra yeniden ele alınıp aslında toksik olmadığı ortaya çıkarılıncaya kadar bu durum sürmüştür. GCP olsaydı ne değişmiş olurdu? GCP'nin dayattığı kurallardan biri şudur: ilacın kuralına göre üretilmiş olması, ilacın gerçek anlamda bir ilaç olduğunun kanıtlanması ve hatta bunun için etik kurul dosyalarına üretim sertifikasını bile koymamız isteniyor. Parasetamolün 1890'larda yapılan o çalışmasında, saflaştırılmamış bir parasetamol kullanılmış ve karaciğer toksisitesine dayanılarak, 60 yıl süreyle bu yanlış nedenden ötürü tıbbın kullanımına girememiştir diyebiliriz.

Klinik çalışmalarla ilgili ilk düzenleme 2. Dünya Savaşı'ndan sonra yapılmıştır. İkinci Dünya Savaşı yıllarında Nazilerin ve Japonların savaş esirlerine ve toplama kamplarındaki çeşitli insanlara yaptıkları medikal testler ve deneyler (örneğin soğuğa dayanma, belli koşullarda ölüm süresinin araştırılması gibi), işkence kabul edilebilecek medikal araştırmalar ilk tepkileri doğurmuş ve savaşın hemen sonrasında Nürnberg Mahkemeleri sırasında ilk kuralların ortaya konmasıyla Nürnberg Yasası ortaya çıkmıştır. Nürnberg Yasası çok iyi niyetli ancak çok naif bir yasadır ve çok fazla bir düzenleme getirmemiştir. Nürnberg Yasası'nın hükümlerine göre insanlar bir çalışmaya alınacaksa, öncelikle rızaları alınmalıdır ve bu çalışmadan zarar görmemeleri bir önkoşuldur. Özetle söylediği budur. Ancak Nürnberg Yasası uygulamaya geçmemiştir. Aradan yıllar geçtikten sonra iki olay, bir felaket ve bir rezalet, dünyanın aklını başına getirmiştir.

Bunlardan bir tanesi Amerika Birleşik Devletleri'nde devletin yaptığı bir çalışmadır. *National Institutions of Health*'in yaptığı Tuskegee Sifilis

Çalışması denilen bir çalışmada, fakir eyaletlerden birinde bir tarım kasabasında yaklaşık 400 siyah tarım işçisi bir sifilis çalışmasına alınmış, bir gruba penisilin verilirken, diğer gruba ise plasebo verilmiştir. Araştırıcılar, tüm hastalara ilaç verdiklerini söyleyerek, yani insanları kandırarak çalışmada tutmuş ve bir grup hasta yaklaşık olarak 20 yıl süreyle tedavi edildiklerini zannederken tedavisiz bırakılıp hastalığın ilerlemesine izin verilmiştir. Sonra çalışmada yapılmakta olanların farkına varılmıştır.

İkinci olay, birinci olayla aşağı yukarı eşzamanlı olarak gerçekleşen ve hepimizin çok daha yakından bildiği talidomid faciasıdır ve 1960'ların başında yaşanmıştır. Bu iki olay dünyanın aklını başına getirmiş ve tüm dünyada koordinasyonlu olmayan, birbirinden bağımsız düzenleme çalışmaları başlamıştır. İlk uyanan ve harekete geçen Amerikalılar olmuştur. O yüzden GCP'nin doğum yeri Amerika diyoruz. Bu iki olaydan sonra 1967'ye doğru komisyonlar toplanıyor, 1970'lerde FDA klinik araştırma yapanlar için ilk kılavuzunu ve ardından da yaklaşık olarak 50 tane kılavuzu çıkartıyor. Her alan için, romatoloji çalışmaları için, kardiyoloji çalışmaları için, çeşitli alanlar için ayrı ayrı olmak üzere çok sayıda kılavuz üretiliyor. Özellikle yeni ilaç çalışmaları, yani henüz ruhsatlanmamış yeni moleküller için, inanılmaz katı kuralları olan IND çalışma kurallarını getiriyor. Bu noktadan sonra da o kurallara uymayan çalışmaları veri olarak kabul etmiyor. Dolayısıyla başka ülkelerden gelen çalışmaları düşük standartlarda oldukları gerekçesiyle reddediyor ve bir ilaç şirketi Amerika Birleşik Devletleri'nde bir ilacı ruhsatlandırıp, pazarlamak istiyor ise Amerika'da, Amerikalı'ların kurallarıyla çalışmalarını yapıp ruhsatını almak zorunda kalıyor.

Eşzamanlı olarak başka ülkeler de süreçleri inceleyerek özeleştiri yapıyorlar ve ulusal GCP'ler ortaya çıkmaya başlıyor. 1990'larda Avrupa GCP kuralları ortaya çıkartılıyor ve Avrupa'daki ilaç şirketleri de bunlara uyarak çalışmaya başlıyorlar. Dünya ilaç pazarına baktığınızda Japonya, Amerika Birleşik Devletleri ve Avrupa'dan sonra 3. sırada yer almaktadır. Japonlar da kendi GCP kurallarını ortaya koyup 1991'de yürürlüğe sokuyorlar. Dünya öyle bir noktaya geliyor ki, Amerika Birleşik Devletleri, Avrupa ve Japonya farklı farklı GCP'lere göre farklı farklı çalışmalar yapıyorlar ve tümü bir diğerinin verilerini kabul etmeme eğilimi gösteriyor. Dolayısıyla bir ilaç şirketi dünyanın tamamında bir ürünü ruhsatlandırıp satmak istiyorsa aynı çalışmayı üç farklı coğrafi lokasyonda, üç ayrı kez, üç farklı kurallar dizgesine uyarak tekrar etmek zorunda kalıyor. Diyelim ki, bir ilacın ruhsatlandırması için yapılması gereken çalışmalar 100 milyon dolara mal olacaksa (rakamı varsayıyorum), bu yüzden 300 milyon dolara mal oluyor ve tekrarlamalar, her ne kadar paralel yapmaya çalışırsanız çalışın, gereken süreyi uzatıyor. Dolayısıyla bunların bir çatı altında birleştirilmesi bir hedef olarak görünüyor; bu sorunu ortadan kaldırmak ve dünyanın uyacağı bir tek GCP ortaya çıkarmak için yola çıkılıyor.

Sonuç olarak ICH GCP'nin ortaya çıkışının altındaki nedenin temelde para olduğunu söyleyebiliriz. *International* olarak nitelendirilen ICH GCP toplantılarına bu üç bölgeden temsilciler katılıyor. İlaç şirketleri ve ruhsatlandırma otoritelerinin temsilcilerinin yanısıra, gözlemciler de bu toplantılara katılıyor. Türkiye'ye pasif gözlemci bulundurma hakkı tanınıyor ve Türkiye adına bir kişi gidiyor. ICH GCP 1996'da finalize olup bütün dünyada kullanılmaya başlanıyor. Amerika Birleşik Devletleri ICH GCP'yi bazı çekincelerle kabul ediyor ama sonuçta ABD de kabul ediyor. Şu anda bütün dünyada ICH GCP, bütün klinik çalışmalarda uyulan kurallar dizgesi olarak kullanılıyor.

Türkiye'de bu süreçte durum şu düzeydedir: Türkiye başlangıçta beklenmediği kadar hızlı davranıyor. 1991'de Avrupa GCP'si yayınlandıktan hemen sonra 1993'te, klinik araştırmalar yönetmeliği ve ardından 1995'te İKU (İyi Klinik Uygulamalar) kılavuzu yayınlanıyor. Bunlar temelde, Avrupa GCP'si kılavuz alınarak gerçekleştirilen yasal düzenleyicilerdir ve Türkiye, klinik araştırmaları düzenleyen kuralları açıkçası birçok Avrupa ülkesinden çok daha erken olarak yasalaştırmış oluyor. Ne yazık ki sonrası gelmiyor ve o noktada kalınıyor. Türkiye'nin yasal düzenlemelerini yapmasından sonra birçok şey değişiyor. ICH GCP 1997'de yürürlüğe giriyor, önceki doksandan fazla ulusal GCP kılavuzu artık geçersiz diye kaldırılıyor. Türkiye ise yürürlüğe soktuğu yasal düzenlemelerin revizyonunu 1998'den beri hala yapmaya çalışıyor ancak 1993'teki Klinik Araştırma Yönetmeliği ve 1995'teki İyi Klinik Uygulamalar kılavuzu bizim

kabul ettiğimiz ve klinik çalışma yaparken, klinik çalışmalarımızı uydurmamız gereken yasal düzenleyiciler olarak kalıyor.

Bu anlattıklarım klinik çalışmalarda etik düzenlenlemelerin tarihçesiydi. Kim araştırıcı olabilir diye baktığımızda bunun da belirli kurallarının olduğunu görüyoruz. Öncelikle formel bir eğitimden geçmiş olmak gerekiyor. Bir araştırıcının doktor olması, uzmanlığının olması, uzmanlık alanında 5 yıldır çalışıyor olması, incelenen, araştırılan alanda deneyim sahibi olması ve bunların hepsinin belgelenmiş olması gerekiyor. Araştırıcı olmak istiyorsanız, çalışmaya ayıracak zamanınızın olması gerekiyor. Yapılacak çalışmada hedeflenen ve araştırılacak olan noktaların sizin bilimsel ilgi alanlarınızın içinde yer alıyor olması, hedeflenen hasta sayılarına ulaşabileceğinizden emin olmanız ve aynı hasta grubu için yarışan başka çalışmalarınızın olmaması gerekiyor. Buraya kadar saydıklarımın hepsi tavsiye niteliğinde olmakla birlikte, gerçekten eksiksiz bir çalışma yapmak ancak bunlara uymakla mümkündür. Şimdi söyleyeceğimse bir dayatmadır: İKU Kılavuzu diyor ki, bir çalışmaya adım atarken o kılavuzun yani Türkiye'deki yasal düzenlemenin sizin omuzlarınıza yüklediği yükün farkında olmanız gerekir.

KATKILAR

Bir Katılımcı: Sondan bir önceki maddede geçen aynı hasta grubu için yarışan iki araştırma yapamazsınız demek aslında aynı anda iki araştırma yapamazsınız anlamına geliyor.

Cem Özesen: Tam bir fizibilite sorusu bu. Eğer A endikasyonunda siz her ay 5 hasta görüyorsanız ve size bir çalışma gelmişse ve o 5 hastanın 5'ini de o çalışmada kullanıyorsanız aynı hasta grubunu hedefleyen ikinci bir çalışma almanız ahlaki olmaz.

Hasan Yazıcı: Aynı noktaya ben de çok katılıyorum. Dr. Özesen "çalışma gelmişse..." dedi; araştırıcıya çalışma gelmez, araştırıcı çalışmayı teklif eder ve onu yapar. Bütün *Good Clinical Practice* başkalarından gelen çalışmaları yapan araştırıcı içinse vay halimize! Gayet açık olarak söylüyorum, burada çok büyük bir sorun var ve katılımcının yerden göğe kadar hakkı var. Araştırıcı varsayımı ortaya koyacak ve düzgün bir araştırıcıysa ondan sonra onu yapacak. Eğer durum böyle değil de, sürekli başkalarının araştırmasını yapma durumunda kalıyor ise, bu fevkalade kötü bir durumdur.

Cem Özesen: İdeal koşulların Profesör Yazıcı'nın belirttiği gibi olduğunu biliyoruz ancak Türkiye'de zaten genel ilaç çalışmalarında araştırıcıların böyle bir rolü yok. Bunun yalnızca şöyle bir pratik sonucu da var; araştırıcılar gerçekten taahhüt ettikleri hasta sayısının çok altında hasta bulabiliyorlar. Örneğin 10 hasta taahhüt ediyor, oysa ancak 2 hasta çalışmaya giriyor ve aynı anda iki çalışmaya katılıyorsa araştırıcı açısından bu zaten yapılmaması gereken bir durumdur. Sizinkinden tamamen farklı bir noktanın altını çiziyorum ancak kastettiğim budur .

Hasan Yazıcı: Bu kılavuz Türkiye için mi yapılmış? Bu Türkiye kılavuzu mu?

Cem Özesen: Evet.

Hasan Yazıcı: Bunu Türkiye için kim yaptı? Türkiye'de bilimselliğin yayılmaması için yapılmış gibi duruyor. Bu olası bir durum değildir; bir grup kendi araştırmaları için kendi planını yapacak ve ardından araştırmalarını gerçekleştirecektir. Türkiye sadece başkalarının araştırmalarının yapılacağı, yalnızca bu nedenle *Good Clinical Practice*'e uyulan bir yer olamaz!

Cem Özesen: Son maddeyi söylüyorsanız; açıkçası birkaç çalışmanız varsa ve kendinize belirli bir zaman hedefi koymuşsanız, bu durum hiçbir çalışmada başarılı olamamanız anlamına gelebilir çünkü hiçbir çalışmaya yetecek kadar hasta bulmanız mümkün olmaz.

Hasan Yazıcı: Benim grubum bu şekilde şimdiye kadar 15 çalışma gerçekleştirdi.

Cem Özesen: İlaç çalışması olduklarını varsayarak söyledim. Uygun hastaları tarayabilmeniz, *monitoring*'le ilgilenebilmeniz, araştırıcılar toplantısına katılabilmeniz, çalışmayla ilişkili prosedürleri yerine getirebilmeniz ve *audit*'ler için zamanınızın olması gerekir. Bir çalışma prensip olarak takım çalışmasıdır; yardımcı araştırıcı kullanmak mümkündür. Türkiye'de çok kullanılmayan bir kavram çalışma koordinatörüdür. Hekim kalifikasyonlarına, yardımcı araştırıcı kalifikasyonlarına sahip olmayan insanları en azından dokümantasyon için kullanmak mümkündür.

Ancak bu insanların mutlaka bir CV verebiliyor olması lazımdır. Yardımcı araştırıcının CV'sinin de araştırmaya uygun olduğunu gösteriyor olması ve klinik araştırma ortamında o çalışmanın yapılması için gerekli olan ekipmanın mevcut olması gereklidir. Baş araştırıcının yapması gereken temel şeyler çalışmanın protokolünü okumak, tartışmak ve ondan sonra kabul etmektir.

Hasan Yazıcı: Bu kabul edilebilir bir durum değildir. Bu dayatmanın bilimselliğin karşısında olduğunu düşünüyorum. Özür dilerim, bu durumun sizden kaynaklanmadığını biliyorum ancak bu üniversitenin ayağa kalkmasına neden olmalı. "Araştırıcı protokolü kabul etmek zorundadır," diye bir dayatma olamaz.

Cem Özesen: Burada anlattığım noktalar tam bir araştırma meselesi değildir; ilaç geliştirme ve ruhsatlandırması için yapılan çalışmaların hepsinin belirli bir harmonizasyonda olması gerekiyor.

Araştırmacının çalışmayı ve çalışma ilaçlarını iyi biliyor olması kurallardan bir tanesidir ve etik kurul onayı almak da yine araştırıcının sorumluluğundadır. Araştırıcının kliniğinin hasta alma kapasitesini iyi öngörebiliyor olması gerekir. Sorumlulukları arasında çalışmaya dahil ettiği hastalardan bilgilendirilmiş olur almak ve protokol tarafından tanımlanmış değerlendirmeleri yapmak ya da yaptırmak, çalışma ilaçlarını saklamak ve sayımını gerçekleştirmek, kendi çalışma takımına bilgi akışını sağlamak, veri toplama enstrümanı olan vaka rapor formlarını doğru doldurmak ya da eğer yardımcıları dolduruyorlarsa doğru veri kaydı olduğunu kontrol ederek teyit etmek, monitorizasyona ve monitöre zaman ayırmak ve genel anlamda çalışmanın sorumluluğunun getirdiği yükümlülükleri üstlenmek de bulunmaktadır. Yine Türkiye'ye özgü bir kural da, araştırıcının yürüttüğü çalışmayla ilgili emniyet verilerini kendi etik kuruluna ve etik kurul üzerinden de Sağlık Bakanlığı'na 6 ayda bir bildirmesi gerekliliğidir.

Protokole uyularak yürütülmekte olan bir araştırmada, çalışmaya dahil edilen hastalarda ortaya çıkan hasardan araştırıcı doğal olarak sorumlu değildir. Ortaya çıkan bu hasarın medikal olarak tedavisi ve/veya tazmini sponsorun sorumluluğundadır. Bunun istisnası araştırıcının protokole uymaması durumudur. Protokole uyulmadığı zaman, sigorta taaddütnamesi işlevini kaybetmektedir.

Türkiye'de henüz araştırma hastalarının sponsoru veya araştırıcıyı dava etmeleri durumu yaşanmadı. Evet, Türkiye'de henüz bunun örneği yaşanmadı ancak böyle bir durumda hastanın hekimi/araştırıcıyı dava etme hakkı vardır. Bir ihmal veya protokole uyumsuzluk varsa, o zaman sigorta şemsiyesi de kalmıyor. Ancak açıkçası şimdiye kadar bunun örneğini bu ülkede duymadım.

Bir Katılımcı: Sponsorun olmadığı, Prof. Yazıcı'nın söylediği şekilde kendi başlattığım ilaç çalışmalarında bu madde nasıl karşılanır?

Cem Özesen: Siz sponsor olarak imzalamıyor musunuz onların altını?

Bir Katılımcı: Sponsor araştırıcı mı oluyor?

Cem Özesen: Evet.

Bir Katılımcı: Araştırıcının kendisi mi sponsor oluyor? Ama ben bunu imzalamam çünkü İngiltere'deki örnek gibi bir hasar oluştuğunda, ben nasıl karşılayacağım?

Bir Katılımcı: Bir yerden bir sigorta bulmanız gerekiyor; bir ilaç şirketinden olabilir veya bir hayır kuruluşundan olabilir.

Cem Özesen: En büyük bağlayıcı şu anda budur. Türkiye'de ilk defa olarak bir sigorta şirketi klinik çalışmalarda sigortalama yapmaya başladı. Bunu da bir haber olarak vereyim.

Yine ilaç çalışmalarında araştırıcının çalışma yaptığı ilaçla ilgili çok kapsamlı bilgiye sahip olması gerekiyor. Sponsor zaten araştırıcıya *investigator's brochure* adı verilen dokümanla çalışma ilacıyla ilgili bilgi veriyor ancak araştırıcının sorumluluğu bu bilgiyi haiz olmaktır. Bu ICH GCP'nin de tanımladığı bir yükümlülüktür. Bir audit bulgusu olarak zaman zaman ortaya çıkmakta ve araştırıcının ilaçla ilgili bilgilerinin yetersiz olduğu durumlarla karşılaşmaktayız.

GCP'yle ilgili olarak üzerinde durmak istediğim önemli bir nokta advers olay bildirimleridir. Advers olay bildirimleri özellikle yeni geliştirilmekte olan ürünlerde, yeni ilaçlarda ilacın emniyetini ortaya koymak için elinizdeki tek enstrü-

man olarak görünüyor. Hele çok hızlanmış ruhsatlandırma süreçlerinin yaşandığı günümüzde, advers olay bildirimleri, ilacın advers olay veri tabanının oluşmasına çok ciddi katkıda bulunuyor.

Hasan Yazıcı: Bu da yanlış bir şey. Dün kontrollü klinik çalışmalarda advers etki veri tabanı olamaz çünkü bu çalışmalara başka grup insanlar girer diye gösterdik.

Good Clinical Practice observational çalışmalarla ilgili değildir. Şu an ilaç çalışmalarından bahsediyoruz ve bahsettiğimiz çalışmaların hepsi ilaç çalışmalarıdır. İlaç çalışmalarını anlattık ve bizim olgularımızın ancak %6'sı bunlara uyuyor dedik: dünyanın her yerinde bu durum böyledir. Koksartroz çalışmalarında, hipertansiyon yan etki sıklığı % 3'tür, oysa osteoartrozda hipertansiyon sıklığı % 40'tır. Ben ne kadar teslim olduğumuzu şu an anlıyorum. kendi iflasımızı biz hazırlamışız. Bunlar "hekimler araştırıcı olamayacaklar" anlamına gelir. Çok açık söylüyorum bu durum, "hekimler ve üniversite öğretim üyeleri devamlı başkalarının hazırladıkları çalışmaları yapmalılar," anlamına geliyor. "Yan etkileri de buradan topluyorum," diyorsunuz. Böylelikle bir firma çıkıp selekoksible 20.000 kişilik yan etki çalışması başlatabiliyor. Yan etki çalışması demek doğrudan doğruya Helsinki Bildirgesi'ne karşı çıkmak demektir; bilimsel olarak yan etki çalışması yapamazsın. Yan etki çalışması yapabilmek için hipertansiyona, konjestif kalp yetmezliğine eğilimi olanları, ülser geçirmiş olanları alman gerekir.

Cem Özesen: Bir düzeltme yapıp ruhsatlandırma öncesi ilaç çalışmaları, yani faz 4 aşamasına gelmemiş çalışmalarda başka advers olay toplama şansımızın olmadığını söylemek isterim.

Hasan Yazıcı: Peki, bize gelen ve etik kurula giren çalışmaların yüzde kaçı bu şekildedir? Cevabı % 0.5'tir. Prof. Direskeneli ileride söyleyecektir; çok az sayıda faz 3 var, faz 2 zaten hiç yok.

Bir Katılımcı: Prof. Yazıcı'nın da az evvel dile getirdiği gibi, etik kurulların İyi Klinik Uygulamaları çerçevesinde çalışması şu aşamaya gelmiştir: etik kurullar bu kuralların denetlendiği ve uygulandığı yerler olarak varlıklarını sürdürüyorlar ve maalesef kurumsal bir değerlendirme yapmaktan ziyade tamamen İKU kurallarının uygulandığı kurullar olarak çalışıyorlar. Bunun da şöyle bir sıkıntısını görüyoruz: etik kurullar maalesef az evvel sıkıntılarını tartıştığımız bir takım maddeleri ilaç çalışması dışındaki çalışmalarda da doğrudan uygular haldeler. İlaç çalışmaları dışında da artık aynen ilaç çalışması mantığı içerisindeler; araştırmacıya bakışı, araştırma kurallarını değerlendirmesi, hatta ve hatta sigortalamayı doğrudan araştırmacıya yüklemesiyle etik kurullar artık inanılmaz bir düzende çalışıyorlar. Gerçi Sağlık Bakanlığı ilaç dışı çalışmaların kendisine raporlanmasını istemiyor, böyle bir düzenleme yok, ama maalesef etik kurulların işleyişi bu mantıkla yürüyor. Bunu nasıl çözmek mümkün?

Haner Direskeneli: Bu etik kurulların deneyimli olmasıyla, etik kurullarda klinik araştırmacının bulunmasıyla, birazdan söyleyeceğim gibi her dekanla değişmeyen etik kurulların olmasıyla mümkündür. Tabii sizin dediğiniz türde sorunların etik kurullar tarafından algılanabilmesi için araştırıcıların böyle çalışmalar getirmesi de gerekir. Bırakın işin daha sonraki basamaklarını, bana etik kurullarda gelen çalışmaların kaç tanesinin birinci basamaktaki bilimsel standartları karşılıyor olduğunu gördüğünüzde, işin zor olduğunu algılıyorsunuz. Bu nedenle birazdan söyleyeceğim gibi çalışmaların dizaynına bile müdahale etme hakkını kendimde görüyorum çünkü dizayn konusunda temel fikri olmayan araştırıcılar önüme yeşil çayla ilgili araştırma getiriyorlar. Biz Türkiye'de daha o noktadayız, bunu unutmayalım. İlaç çalışmaları yine de biraz daha yüksek bir standardı gerektiriyor ancak tabii dediğiniz türde sorunlarla karşılaşıyoruz.

Hasan Yazıcı: Türkiye'deki sorunun farkındayım. Örneğin, bir firmanın piyasadan kalkan ilacı için 20.000 kişilik bir advers etki çalışması yapma planladığını öğrendim. Toplumun ilaç firmalarına ve bilim adamlarına güveni aşınmaya başladı. Biz bilimsel açıdan ilaç firmalarına hayır demeyi öğrenmedikçe bu güven daha da çok kaybolacaktır.

ETİK KOMİTE GÖRÜŞÜ

Haner Direskeneli

Konuşmama başlarken toplumun ilaç araştırmalarına bakışını etkileyen bir diğer olaydan bahsedeceğim. Dr. Özesen Tuskegee çalışması ve talidomid olayından bahsetti. Bunlardan sonra gelen üçüncü önemli olay, şimdi anlatacağım çalışmadır. www.jointandbone.org sitesinden 16 Mart tarihli bir alıntım olacak; ancak gazetelerde de söz edildi, belki bilenleriniz vardır. Yeni bir ilaç bir Faz I çalışma çerçevesinde, İngiltere'de Northwick Park Hastanesi'nde ilk defa 6 sağlıklı gönüllüye verildi ancak tüm gönüllülerde ağır yan etkiler görüldü. İki tanesi birkaç saat içinde çoğul-organ yetmezliğiyle yoğun bakım tedavisi gerektirir hale geldi. Söz konusu ilaç, bir Alman firması tarafından geliştirilen ve romatoid artrit, multipl skleroz ve KLL için geliştirilmesi planlanan bir T-hücre uyarıcısı antikordur. Son 1 aydır gündemde olan bu konuyla ilgili Türkçe bir gazete haberiyle devam etmek istiyorum:

Radikal: 4 Nisan 2006: 'Fil adam hayata döndü'.

Gazeteler toplumun bakış açılarını belirlediğinden, entelektüel kesime hitap ettiği söylenen bu gazetenin bakış açısı önemlidir. Denek kendisine baş ağrısı ve mide bulantısından başka bir şey hissetmeyeceğinin söylendiğini söylüyor, ilacı aldıktan bir saat sonra şiddetli bir baş ağrısı saplanıyor, nefes darlığı oluşuyor, ağrıdan ağlayacak hale geliyor. Kendisini ziyaret eden arkadaşı onun bir "fil adam" gibi olduğunu, kafasının şiştiğini söylüyor. Olanları Hindistan'daki ailesine anlatamayan genç adam ilaç firmasından alacağı 2000 sterlinlik deney ücretiyle bilgisayar almayı planlıyormuş. Daha önceden iki tane ilaç çalışmasına katılmış; her iki deneyde de bahsedilen yan etkilerden daha fazlasını yaşamadığı için yine yaşamayacağını varsayıyormuş. Kendisine deney öncesinde ilacın hayvanlar üzerinde olumsuz etkisinin olmadığı söylenmiş. Hastaneden yeni çıkmış, diğerleri henüz çıkamamış durumda ve son cümle olarak "bana mutluluğumu ve sağlığımı asla geri veremezler," diyor.

Şimdi bu konuda ne gibi sorunlar var? Birincisi; bu çalışmada temel bir hipotez sorunu var. Bu hem araştırıcıların, hem de etik komitenin sorunudur çünkü bu ilacın bir uyarıcı molekül olmasına karşın hipotez, bu molekülün sadece regülatuar (anti-inflamatuar) hücreler üzerinde immünostimülatör etki yapacağı, diğer hücreleri uyarmayacağı üzerine kurulmuş.

Profesör Doktor, Marmara Üniversitesi, Tıp Fakültesi, Romatoloji Bilim Dalı

Gerçekten de bildirildiği kadarıyla, maymunlarda yapılanlar da dahil olmak üzere, hayvan deneylerinde majör bir yan etki görülmemiş ve bu durum çalışmanın bir sürprizi olarak karşılarına çıkmış. Ancak bu durum bundan sonra benzer Faz 1 çalışmaların daha zor yapılacağının da bir göstergesidir. İkinci olarak çok önemli teknik bir sorun var: 6 kişiye birden aynı anda ilaç verilmiş, hatta birinci hasta komaya girerken ikincisi ilacı almak üzereymiş. Bu tabii son derece önemli bir sorunu temsil ediyor: gönüllüler ve sağlık çalışanları da dahil olmak üzere toplum yan etki olabileceğine inanamıyor. Sorun şu: ilaç çok akut etkili. Hastalara tek tek uygulasalardı, daha doğru olurdu. Evet, biraz bekleseler ve ilacı sırayla uygulasalardı, ilk hastadaki yan etkiyi görebileceklerdi. Tabii en sonunda da finansal ve hukuki sorunlar var. İngiltere'de bu tür çalışmalar için yapılan sigorta 30-40.000 sterlini geçmeyen bir tutarla sınırlıymış. Yani hasta Amerika'da olsa, ilgili ilaç firmasından birkaç milyon dolar sigorta alabilecekken, İngiltere'de alamayacak. Sonuçta da olay Türk basınına da böyle yansımış durumdadır: '2.000 dolar için hayat tehlikeye atılıyor.' Gelinen bu çok zor nokta, bundan sonra bu tür çalışmaların nasıl sürdürüleceği konusunda, ne kadar dikkat edilmesi gerektiğini gösteriyor.

Şimdi bu noktada, Vedat Hamuryudan'ın konuşmasında tartıştığı plasebo kontrollü çalışmalar konusuna döneceğim. Çünkü etik kurullarla ilgili olarak en çok tartışılan noktalardan biri budur. Dr. Hamuryudan'ın da bahsettiği gibi, Helsinki Deklarasyonu'nun orijinal versiyonuna göre, kontrollü çalışmalar dahil olmak üzere, her tıbbi çalışmada her hastanın kanıtlanmış en iyi tedaviye ulaşma hakkı mevcuttur. Bunu kabaca yorumlarsak, kontrollü çalışmalar ve özellikle plasebo kullanımı uygun değil gibi görünüyor. Ancak bunun çok önemli bir sorun olduğunu birazdan tartışacağım; bu çerçevede Helsinki Deklarasyonu'nun 2002'deki düzeltilmiş son versiyonunda yeni bir tanımlama mevcuttur: "Yeni bir yöntemin yararları, riskleri ve etkinliği kullanılmakta olan en iyi koruyucu, tanısal veya tedavi amaçlı yönteme karşı test edilmeli," deniyor. "Kanıtlanmış koruyucu, tanısal veya tedavi amaçlı yöntemin bulunmadığı çalışmalarda plasebo kullanımını veya tedavisiz izlemi dışlamamaktadır," deniyor ve bu çerçevede de bir açıklama notu var. Sürekli imzaladığımız şu formların son sayfasında yer alıyor: "Dünya Tıp Birliği genel olarak plasebo kontrollü çalışmaların yalnızca kanıtlanmış tedavi yokluğunda kullanılabileceğini onaylar." Ancak kanıtlanmış bir tedavi bulunsa da, aşağıdaki koşullarda, plasebo kontrollü deney yapılabileceği de eklenmiştir. Bu koşullardan biri, zorunlu ya da bilimsel olarak geçerli metodolojik nedenler dolayısıyla farklı bir yöntemin etkinliğinin araştırılmasıdır. Bu bilimsel açıdan tartışmamız gereken plasebo kontrollü çalışmanın gerekliliği kavramını içeriyor.

İkinci bir koşulsa yapılan araştırmanın hafif bir hastalıkta yapılıyor olmasıdır. Tabii burada hafifin tanımı belli değil, kastedilen baş ağrısı, dispepsi veya benzeri başka sorunlar olabilir. Önemli olan, plasebo alanların herhangi bir ek ciddi veya geri dönüşümsüz zarar riskiyle karşılaşmayacak olmasıdır. Bu durumda da plasebo kontrollü çalışma yapılabiliyor. Böylelikle bu konuda bir esneklik sağlanmış durumdadır. Buradaki ikinci soru kanıtlanmış en iyi tedavi yönteminin ne olduğu sorusudur. Bunu çok kabaca yorumlarsak şunu diyebiliriz; aktif kontrollü çalışmayı da yapamıyor olmamız gerekir çünkü ortalıkta bir kanıtlanmış tedavi varsa yeni yaptığımız her tedavi kanıtlanmamış olur ve deneysel nitelik kazanır. Tabii bu yanlış bir yorumdur çünkü aktif kontrollü, eşdeğerlilik çalışmaları her zaman plasebo kontrollü çalışmaların yerini alamıyor. Bunun nedeni, bu çalışmaların öncelikle şöyle bir temel problemi olmasıdır: aktif kontrollü çalışmadaki standart tedavinin plaseboya üstün olduğunu varsayıyoruz, ama tarihi kontrollerle kıyaslayarak. Böyle olunca, çalışmanın içinde plasebo olmadığı durumda her zaman için kontrole göre daha etkin olduğunu varsayıyoruz.

FDA'den iki araştırıcının bu konuyu tartışan iki derlemesi çok yararlı yazılardır: Birçok alanda efektif kabul edilen tedavilerin % 30 ila 50'sinin plaseboya üstün olmadığı gösterilmiş, meta analizlerde üstün çıksa bile tek tek bakıldığı zaman birçok çalışma negatif çıkabiliyor. Yani bu anlamda çalışmanın içinde plasebo kolu kullanmazsanız, o tedavinin etkinliğini değerlendirmeniz zor olabilir. Bu yüzden ideal çalışma yeni tedavi, standart tedavi ve plasebo kollarını içermelidir. Böyle planlanmış bir çalışma hem etkisiz bir yeni tedaviyi gösterebilir (standart tedavi plaseboya üstündür ama yeni tedavi plaseboya üstün değildir); hem de etkisiz bir çalışmayı göst-

terebilir (ne yeni tedavi, ne de standart tedavi plasebodan üstün değildir). İdeal olarak karşılaştırmalı çalışmaların bu koşullarda yapılması gerekir.

Peki, diğer yaklaşımlar nasıl olabilir? Yeni ajan standart tedaviye ek tedavi (*add-on*) olarak kullanılabilir: örneğin metotreksata dirençli hastalara metotreksat + başka bir tedavi vermeniz gibi. Ancak monoterapi olarak etkinliği araştırmada bu, yetersiz kalan bir dizayn olur. Öncelikle ilacın tek başına kullanımını değerlendiremezsiniz. İkinci olarak, yeni ve standart tedaviler farmakolojik olarak birbirine çok benzer moleküllerse bu dizayn uygun olmayabilir. Anti-TNF ajanların etki mekanizması metotreksattan çok farklı olduğu için bu yaklaşım uygun olabilir ama farmakolojik yollar birbirine çok benzerse, örneğin metotreksat + azatioprin ile çalışma yapamazsınız. İkinci bir yöntem erken kaçış (*early escape*) modelidir: kısa süreyle plasebo kullanmak ve plasebo kullanımı sırasında hastalığın kötüye gittiği fark edildiğinde, geri dönüşümsüz zarar oluşmadan, hastayı etkin tedavi koluna geçirmek. Son olarak Dr. Hamuryudan'ın JRA çerçevesinde gösterdiği şekilde bir geri çekme (*withdrawal*) çalışması yapmak. Deneysel bir tedaviye başlanır ve eğer tedavi etkinse belli bir noktada ilaç kesilerek plaseboyla karşılaştırılır. Şimdi bu çerçevede etik yönden nasıl bir çalışma tasarımı yapabiliriz? (Şekil 1).

Öncelikle, çalışılan hastalıkla ilgili etkili bir tedavi var mı? Birinci sıradaki sorumuz bu olmalıdır. Yoksa, plasebo kontrollü çalışma yapabilirsiniz. Etkili tedavi varsa o zaman soracağımız soru, aktif kontrollü eşdeğerlik çalışmasının yeterli olup olmayacağıdır? Yeterliyse bunu yapabilirsiniz. Yetersizse o zaman şunu sormak gerekiyor: Tedavi sağkalım ya da geri dönüşümsüz mortalite üzerine etkili mi? Değilse, kısa süreli bir baş ağrısı çalışması gibi veya bizim alanımızdan örnek verecek olursak belki bir ostreoartroz ya da fibromiyalji çalışması gibi plasebo kontrollü çalışmayı yine yapabilirsiniz. Diyelim ki onkoloji alanında tedavi etkiliyse, o zaman ekleme tedavisi yapabilir mi ve bu model uygun mu sorusunu sormak gerekir. Bu model de uygun değilse, etik açıdan uygun kısa dönem plasebo kontrollü çalışma yapabilir misiniz? Diyelim bir aylık kısa bir plasebo kontrollü çalışma uygun mudur? Bütün bunlar uygun değilse, o zaman başa dönmek gerekir. Etkili tedavi dediğiniz tedavi gerçekten kanıtlanmış mı? Siz etkili tedavi olduğunu düşünüyor olabilirsiniz ancak plasebo kontrollü çalışmalar ve meta analizler göstermiyorsa (ki dün gördük ki birçok alanda meta analiz yapınca etkinlikler ortadan kalkıyor) o zaman etik komiteyi ikna ederek hala plasebo kontrollü çalışma yapmaya hakkınız var ancak böyle değilse o zaman artık tek çözümünüz kalıyor: yeni tedaviyi aktif kontrolle karşılaştırmak. Tabii bu ça-

Şekil 1: Etik Yönden Araştırma Tasarımı
(Ellenberg, Ann Intern Med 2000; 464)

Etkili tedavi var mı?	(-)	Plasebo kontrollü çalışma (PKÇ)
Aktif kontrollü eşdeğerlik çalışması yeterli mi ?	(+)	Aktif kontrollü eşdeğerlik çalışması
Tedavi sağkalım / geri-dönüşümsüz morbidite üzerine etkili mi?	(-)	Plasebo kontrollü çalışma
Ekleme (add-on) çalışma yeterli mi ?	(+)	Ekleme çalışması
Etik açıdan uygun kısa-dönem PKÇ yeterli mi ?	(+)	Kısa dönem plasebo kontrollü çalışma
Etkili tedavi "standart tedavi" olarak kabul ediliyor mu ?	(-)	Plasebo kontrollü çalışma
Yeni tedavi aktif kontrole üstün olabilir mi?	(+)	Aktif kontrollü üstünlük çalışması

lışmalar da, en büyük örneklem sayılarının kullanılmasını gerektiren üstünlük çalışmalarıdır.

KATKILAR

Hasan Yazıcı: Burada da bir eksik var. Diyelim ki, "Behçet hastalığında merkezi sinir sistemi tutulmasının etkili tedavisi var mı?" diye sorduk kendimize. Bu sorunun yanıtı olumsuz. Plasebo kontrollü çalışma yapabilir miyiz? Çok kısıtlı bir durumda yapabiliriz, o da ancak araya başka bir *escape drug*'la girersek.

Haner Direskeneli: Niye etkili tedavinin olmadığını düşünüyorsunuz? Burada etkili tedaviden kastettiğim sizin sadece etkili olduğuna inanmanız; steroidin etkili olduğuna inanıyorsunuz.

Hasan Yazıcı: Hiçbir algoritma bu dünyada bir deneyimli hekimin önüne geçemez. Hiçbir algoritma, bir deneyimli hekim ve hekim grubunun bir araya gelip, düşünüp, bilgilerini vicdanlarıyla birleştirip karar vermelerinin üstünde görülemez. Endüstrinin amacı ayrıdır ve para kazanmaktır. Ticaret, ahlak getirir ve onun karşısında, onu dengede tutacak, vicdanla bilgiyi birleştiren bir grup yok ise durum çok vahimdir. Şu andaki durumumuz ikinci duruma benziyor, benim gördüğüm budur.

Haner Direskeneli: Ama benim söylediklerim tam dediğiniz noktayla örtüşüyor. Başlangıçta zaten bir grup araştırıcı, bu çalışmayı dizayn edenler bir araya geliyorlar; etkili tedavi olup olmadığına bakıyorlar. Bu artık endüstri süreci değil, tamamen araştırıcının sürecidir.

Buradan itibaren, daha az tartışmalı konulara geçiyorum: İlaç araştırma dönemleri. Kısaca hatırlarsak, birinci dönem (Faz 1) biraz önce aldığımız örneğin bulunduğu dönem, yani sağlıklı gönüllülerde deneme sürecidir. İkinci dönem (Faz 2) az sayıda seçilmiş hasta üzerinde doz sınırlarının, aktivite ile yararın denenmesi, üçüncü dönem (Faz 3) daha geniş sayıda hasta üzerinde incelenmesi ve dördüncü dönem (Faz 4) ise farmasötik müstahzar olarak ruhsatlandırılmasından sonraki dönemdir. Bu dönemleri şu anda bizim bağlı olduğumuz hukuki çerçeveyi oluşturan 1993 tarihli İlaç Araştırmaları Yönetmeliği'ne göre tanımlıyorum.

İlaç araştırmalarında taraflar kimlerdir? Denekler, araştırıcılar, araştırma kurumları, düzenleyici devlet kurumu, endüstri ve klinik araştırma kurumları (CRO) gibi aracı kurumlar tarafları oluşturur. Biz etik komite olarak, şu an itibarıyla, araştırma kurumu ile düzenleyici devlet kurumu arasında yer alıyoruz. Peki etik kurulun varoluş amacı nedir? Etik kurulun temel amacı, biyomedikal araştırmalara katılan deneklerin onurunu, hakkını, güvenliğini ve esenliğini korumaya hizmet etmektir. Araştırmanın amaçları bunun önüne geçemez. Öncelikle bireylerin onuruna saygıyı korumalıyız. Tabii bu çerçevede de bir adalet ilkesi vardır; sonuçlarından toplumun tüm kesimlerinin yararlanabileceği çalışmaları yaparken fayda ve zararları tüm kesimlere eşit dağıtmak gibi bir yükümlülüğümüz mevcuttur. Bu WHO'nun etik komiteler için önerdiği bir genel yönlendirmedir. Bu çerçevede etik komite bir çalışmanın yalnızca etik değil, tüm yönlerini değerlendirmek durumundadır. Bu etik kurula sık sık sorulan bir sorudur. Etik komite çalışmanın dizaynına karışabilir mi? Evet karışabilir çünkü etik komite olayın bilimsel, hukuki, ekonomik vb. tüm yönlerini değerlendirmekle yükümlüdür. Bu çerçevede etik komitelerin hukuki yapısı da önemlidir. Etik komiteler başlangıçtaki kendiliğinden bir araya gelip değerlendirme yapan komiteler olma niteliğinden çıkarak, yavaş yavaş yönetmelikleri olan, çalışma düzenini oluşturan, hem deneğe, hem de diğer taraflara karşı hukuki yapısını oluşturmaya çalışan bir yapıda gelişmektedirler. Etik komitenin rolü ne olabilir? Şu anda yalnızca danışmanlık yapma noktasındayız. Aslında etik komite olarak, karar verici olan üniversite, devlet vb. gibi kurumsal oluşumlara danışmanlık yapıyoruz. Ama araştırmalarla ilgili eğitim, süpervizyon vb. gibi konular için de bir ana kurul olabiliriz ya da şu anda çok az yapılan, çalışmaların denetlenmesi sürecinde doğrudan yer alabiliriz. Şu an için biliyoruz ki; Türkiye'de klasik anlamda ilaç çalışmalarında sadece araştırma kuruluşları denetleme yapıyor; etik komitelerin bunu pek yaptığını görmüyoruz.

Etik komite yalnızca deneğin hakkı için bulunduğundan tüm diğer taraflardan finansman olarak da, denetim olarak da bağımsız olmak zorundadır. WHO genel olarak devletlerin ulusal, kurumsal ve yerel düzeyde bağımsız, çok disipliner, çok sektörcü ve çoğulcu bir yapıda etik komite kurmasını destekliyor. Şimdi bu çerçevede-

ki sorulardan birincisi, etik komiteyi kim oluşturmalı? Şu anki biçimiyle merkezi otorite bir merkezi etik kurul oluşturuyor, araştırma kurumları da sıklıkla olduğu gibi dekanlık ya da başhekimlik tarafından atanan bir yerel etik kurul oluşturuyor. Bir örnek vereyim: Bizim etik komitemiz 2 yıl önce yeni dekanımızın göreve gelmesiyle birlikte ben dışımda bütün üyeleri değiştirilmesiyle yeni baştan oluşturuldu. Bu yanlış bir uygulamadır çünkü etik komitenin kendi oluşturduğu yapıyı, bilgilenme sürecini yok sayma anlamına geliyor. Bağımsız bir süreçte de iki sorun var: Etik kurul tamamen bağımsız olabilir ama bunun riski iç çoğalma (*inbreeding*) olmasıdır, yani etik komite kendi kafa yapısındaki üyeleri sürekli atadığı için farklı görüşlerin yansıması az olabilir. Bu bütün bağımsız kurumlar için geçerli bir durumdur ve tabii değerlendirmeyi etkileyebilir. İkinci konu ise çıkar çatışmasıdır. Etik komite üyelerinin projeleri geldiği zaman etik komite ne kadar bağımsız olabilir?

Şu an itibarıyla yerel etik komitelerin 1993 tarihli yönetmeliğe göre belirlenmiş bir yapısı mevcuttur. Üç klinisyen doktor, birer klinik farmakolog veya tıp doktoru farmakolog, eczacı, biyokimya uzmanı ve patolog bu kurulda minimum olarak yer almalıdır. Bu yönetmeliğin neden bir biyokimya uzmanı ve patologun bulunması gerektiğini istediğini anlamak zordur. Yine yönetmeliğe göre mümkünse bir tıbbi deontoloji uzmanının ve yapılacak çalışmanın özelliğine uygun bir uzmanın bulunması isteniyor. Diğer bir deyişle, her etik komite toplantısına ayrı bir uzmanın çağrılması bekleniyor. Bu ise yönetmeliğin aslında o sırada yeterince düşünülmeden hazırlandığını düşündürüyor. Şu an itibarıyla yerel etik komitelerin görevi 1 ve 4. fazlar arasındaki protokolleri incelemek ve uygun gördüğünü bakanlığa göndermek, 4. fazdaki protokolleri de gene uygulanıp uygulanmayacağına karar verdikten sonra Sağlık Bakanlığı'na bildirmek ve lüzum gördüğünde de araştırmaları durdurmaktır. Etik komitelerin görevleri bu kadarla sınırlanmıştır ancak şimdi yeni bir taslak var. Biraz önce Dr. Özesen'den öğrendiğime göre 1998'den beri devam etmekte olan bir sürecin sonunda en son 2 ay önce önümüze bu taslak konuldu.

Burada bazı ilginç yeni yaklaşımlar var, bunları sizinle paylaşmak istiyorum. Öncelikle artık eğitim ve araştırma kurumları, kamu kurumları ve üniversite hastaneleri dışında A1 grubuna giren özel hastanelerde de araştırma yapılmasına izin verme eğilimi var. Bunun kısmen ilaç endüstrisinin ve belki özel hastanelerin yöneticilerinin (hastaneye herhalde prestij ve para getireceğini düşünerek) istediği bir şey olduğunu düşünüyorum ancak zorluklar içeriyor. Bu çerçevede etik kurulların yapısında da onaylanmış etik kurul diye yeni bir oluşum getirilerek değişiklik düşünülüyor. Burada etik komitenin yetkileri arttırılacak ve merkezi etik komitede olan bir sürü yetki artık yerel etik komiteye geçecektir. Komitenin bileşiminde ise yeni olarak İyi Klinik Uygulamaları (GCP) eğitimi almış (belki endüstri tarafından kafası iyice yıkanmış diyebiliriz) yani ilaç araştırması yapmış klinisyenler ve bunlar dışında, birer hukukçu ve üniversite mezunu bir üye (tanım olarak hasta hakları alanında örgütlenmiş sivil toplum kuruluşlarından gelen ancak sağlık mesleği mensubu olmayan ve klinik araştırmalarla ilgisi bulunmayan yani Batı'daki anlamıyla *layman*) ve bir tıbbi etik uzmanı zorunlu bulunacak.

Etik komitenin görev tanımı biraz daha genişletilmiş durumdadır. Demin 1993 yönetmeliğinde çok sınırlı olarak sunduğum görev tanımı, yeni metinde denetleme dahil araştırma formlarının değerlendirilmesi, sigortaların incelenmesi ve her türlü anlaşmayı değerlendirecek bir formata dönüştürülmüş; yani onaylanmış etik kurulun görevleri artıyor. Peki, onaylanmış etik komite kime karşı sorumlu olacak? Yeni bir Klinik Araştırmalar Danışma Kurulu oluşturuluyor. Merkezi etik komiteyi kaldırıp bunu kurmayı planlıyorlar. Aslında bu kurul doğrudan yetkili de değildir. Sağlık Bakanlığı bu kurula yalnızca görüş soracak. İçinde Barolar Birliği'nden temsilci var, Tabipler Birliği'nden temsilci var, İlahiyat Fakültesi'nden bir temsilci var. Bunun Avrupa Birliği direktifleri çerçevesinde olduğu söyleniyor ve yine üniversite mezunu birileri yer alıyor. Yeni taslak bu şekilde; tabii değişebilir. Peki, klinik ilaç araştırması nerelerde yapılabilir? Tam teşekküllü eğitim hastanelerinde ve üniversitelerde yapılabilir. Şu anda bizi bağlayan yönetmelik bunu söylüyor ve ne yazık ki özel hastanelerin, en iyilerin de bile, ilaç araştırması yapmak mümkün değildir. Gönüllü formuyla ilgili yenilikler var: Gönüllünün kimliğinin gizlenmesi ve iyi bir gönüllü olur formu olması gereklidir, ayrıca gönüllüye ödeme yapılamayacağı belirtiliyor. Tabii farmakokinetik ve biyoeşdeğerlik çalışmaları bu-

nun dışında bırakılıyor. Biraz önce anlattığım İngiltere'deki felaketin yaşanma potansiyeli hala var ancak bunun çözümünün ne olduğunu da bilmiyorum. Faz 1'deki bir ilacı denemeye kimi para vermeden ikna edebilirsiniz? Düşünmek zor!

Araştırmacı kimler olabilir? Şu anki biçimiyle en az 5 yıl süreyle tecrübeli uzman olmak gerekiyor. Yeni taslak bu süre zorunluluğunu kaldırıyor; uzman olmayı yeterli buluyor. Bu da aslında iyi bir yaklaşımdır. Biraz önce konuştuğumuz gibi, bir araştırıcının üzerinde çok fazla çalışma olması durumunda, işlerin nasıl yürüyeceği meselesinde yararlıdır. Şu anki uygulamada bir kurumdaki en kıdemli öğretim üyesi, tüm çalışmalarda sürekli ana araştırıcı görünüyor ancak bazen çalışmanın detaylarından haberli olmuyor. Araştırıcılar bu konuda doğrudan sorumluluk almalı; bu da onu hedefliyor. Bir diğer önemli fark, aslında özellikle yine endüstrinin çok istediği bir şey olan eşzamanlı başvuruya izin verilmesidir. Şu anki uygulamada yerel etik komiteden geçip, merkez etik komiteye başvuru yapılıyor. Yeni taslakta, eşzamanlı başvuru yapılabilecek. Yeni yönetmelik, şimdiki gibi, aslında tamamen destekleyicinin yürüttüğü ama araştırıcının yapıyormuş gibi göründüğü işlerin gerçekten destekleyiciyle yürütülmesini sağlayacak. Şu anki uygulamada sanki araştırıcı araştırma yapıyormuş gibi görünüyor ama aslında sadece bir araştırma elemanı olarak işlev görüyor. İlginç bir nokta da şudur: yeni yönetmelikte genetik çalışmalar için, bunların daha zor ve normal çalışmalara göre daha uzun bir sürede değerlendirileceği düşünülerek, etik komitelere 90 gün süre verilmiştir.

Faz 4 çalışmanın ürün izleme ve klinik araştırma safhası olarak bir tanımı mevcuttur. Yeni taslakta ise ek bir tanımlama kullanıyorlar: girişimsel olmayan (*non-interventional*) çalışma. Burada sorun şudur: Türkiye'de biliyoruz ki, ilaç firmaları etik kurullar önüne, özellikle kendi ilaçlarının kullanılmasını tercih ettikleri ancak adına gözlemsel dedikleri Faz 4 çalışmaları getiriyorlar. Bakanlık bunların aslında gerçek gözlemsel olmadığını, *non-interventional* tanımlamasının burada gerekli olduğunu düşünüyor ve ben de buna katılıyorum. Süreç söyle tanımlanabilir: ürünü ruhsata uygun reçetelendiriyorsun ama reçeteleme hastanın çalışmada yer alma kararından bağımsız gerçekleşiyor. Diğer bir deyişle, polikliniklerde standart hastalara ilaç reçeteleniyor ve o çalışmada da bunlar o dağılımı yansıtıyor. A firmasının gözlemsel çalışması olduğu için hekimin romatoid artrit hastalarının % 90'ı A firmasının ilacını almayacak; birinci tanım budur. Bu bir protokole göre olmayacak, tamamen gerçek kliniği yansıtacak, hastaya ek tanı ve izleme yöntemleri uygulanmayacaktır; standart ne yapılıyorsa o gerçekleştirilecek ve veri analizi için epidemiyolojik yöntemler kullanılacak. Bakanlığın şu anki yaklaşımı bu çalışmalarda sigorta istememe yönündedir ancak bu konuda tam netleşilmiş değildir. Bu tip çalışmalar, bizim post marketing dönem çalışmaları dediğimiz, gerçekten yan etkileri değerlendireceğimiz nitelikteki çalışmalardır. Bu çalışmaları özellikle endüstri desteğinin ağırlıklı olduğu, şu anki biçimiyle Faz 4 denen gözlemsel çalışmalardan ayırt etmek gerekiyor. Bu gerçekten tam klinik pratiğimizi yansıtan *non-interventional trial* tanımıdır.

KATKILAR

Banu Çakır: Benim bir itirazım var. Faz 4 çalışmalar gözlemsel araştırmalar değildir. Bence tam olarak gözlemsel araştırmada, siz yaşama müdahale etmezsiniz, her şey spontandır, yani sadece gözlersiniz. Faz 4 çalışmalarda bir müdahale var ama ilaç hali hazırda ruhsatlanmış. Biraz önce konuşulan çalışmalar grubunda ben de biraz yer aldım; Sağlık Bakanlığı böyle bir çalışma yapmak istiyor çünkü özellikle gözlemsel ya da *post marketing* çalışmalarda amaç ağırlıklı olarak ilaç firmalarının *seeding* yapması: ilaç firması "ben beşbin kişiye gideceğim," diyor. Sadece kendi ilacımı kontrol edeceğim diyor ama Türkiye'de beşbin hasta yok onun kontrol edebileceği ama ilaç firması çalışmayı yapıyor, bu arada bütün hekimler çalışmalardan haberdar oluyor, ilaçtan haberdar oluyor, yazmaya eli alışıyor; ondan sonra üçbininci kişide "ben bu çalışmayı yapamayacağım çünkü beşbin kişiye ulaşamıyorum" deniyor. Zaten Türkiye'de o çalışmayı yapabileceğiniz 5 bin kişi yok ama sonuç ne oluyor? İlaç firması kendi ilacını tanıtmış oluyor. Bakanlık böyle çalışmaları kontrol etmek için bunu yaptı çünkü klinik çalışmalar için bir çalışma grubu vardı ancak bu tür çalışmalar için yoktu.
Bu sefer de şöyle bir sıkıntı oldu: Türkiye'de bir çalışma yaptığınız zaman sponsorun bu çalışmaları ödemesi gerekiyor çünkü Emekli Sandığı artık sizin yaptığınız tam kan sayımlarını ya da fazladan bir kreatinin ölçümünü ödemek istemiyor.

Ancak bu kez ilaç firması da diyor ki, "bu spontan reçeteleme yani biz hastaya bir şey yapmayacağız, sadece onun değerlerini ölçeceğiz". Bu çalışmaları girişimselden ayırmak için girişimsel olmayan denmiş ancak bunlar ruhsat sonrası yaptığınız gözlemsel araştırmalardır. Bu araştırma olacaksa fazladan bir kan sayımının ödenip ödenmeyeceği konuşulmalıdır. Burada tabii geri ödeme kurumlarının da gündeme gelmesi lazım; yani geri ödeme kurumu ne kadarını ödesin? Eskiden her şeyi onlar ödüyordu.

Haner Direskeneli: İlaç dışı değil belki ama endüstri destekli olmayan diyebileceğimiz çalışmaların da yapılmasının gerekli olduğu ve bu durumda da bazen İyi Klinik Uygulamaları ile ilgili tanımların kapsam dışı bırakılabileceği yönünde düşünceler var. Ülkemiz açısından ilaç araştırmalarında ana sorunlar nelerdir? İlk olarak, ilaç firmaları açısından biraz önce anlatılan İyi Klinik Uygulamaları'na uygun klinik araştırma merkezi gereksinimi dünyada artıyor. Basit bir örnekle, artık Amerika'da ağır RA'lı hastaların büyük bir kısmı anti-TNF ajanlar ile tedavi olabildiğinden yeni bir biyolojik ajan geliştirmek isteyen firma yeterince hasta bulamıyor. Bazen sadece bu nedenle bile sosyal güvenliğin daha zayıf olduğu ülkelerde araştırma merkezi arıyorlar; Türkiye de böyle bir ülke olarak görülüyor. Ancak unutmamamız gereken, herkesin bu konuda ciddi sorumlulukları olduğudur. Bence şu anda, en azından İstanbul için denebilir, araştırıcıların bu tür çalışmalar için motivasyonu düşüktür. Motivasyon olmamasında neden belki bu çalışmaların akademik yaşama fazla katkısı olmadığı gerçeğidir. Çünkü çalışma planlaması, analiz, yazım aşamalarında hemen hemen hiçbir katılım olmuyor. O açıdan ciddi bir akademisyenin bu çalışmalarla motive olması akademik yönden mümkün değildir ancak aslında zaten motivasyon başka yerde aranmalıdır. Bütçesi olan bu çalışma sayesinde, o bütçeyle bölüme cihaz ve altyapı alınıyor; belki bir miktar da araştırıcıya geri dönüyor. Araştırıcıya dönen miktar aslında yakın zamana dek daha yüksekti, ama artık bakanlık da bu işi ciddiye alarak "bütün bütçe döner sermayeye yatacak," diyor ve sonuçta döner sermayeden bunu iyi organize eden kurumlar, örneğin sanıyorum Dokuz Eylül, daha fazla araştırma talep ediyor. İstanbul ve Marmara Üniversiteleri gibi kurumlarsa araştırmacıların artık bu işi yapmak istemediği yerler durumundadır.

Tabii ikinci bir sorun bu araştırmaların getirdiği iş yüküdür. Bu iş yükü yardımcı personel gerektiriyor. Araştırıcılar ise sıklıkla yardımcı personel kullanmadan, yeterli zaman ayırmadan çalışmaları yapmak istiyorlar ve o zaman da işler yürümüyor. Belki özelleşmiş ilaç araştırma merkezleri gibi merkezler, diğer bir deyişle ek bir yerde birkaç farklı displinden araştırmanın birlikte yürütüldüğü ve bu işin adının konarak, yalnızca bu amaca yönelik çalışmaların yürütüldüğü merkezler olması bu işi kolaylaştırabilir.

Etik komitelerin ne sorunları var? Şu anda Türkiye'de 82 tane yerel etik komite var. Tabii bunların hepsinin standartları aynı değildir ve Sağlık Bakanlığı bunları azaltmak istiyor. Onaylanmış etik komitelerin amacı kısmen bu olmakla beraber, başka bir sorun etik kurulların buna hazır olmayışıdır. Bizimki gibi bazı yerel etik komiteler hala zamanlarının önemli bir kısmını lokal dosyalarla geçiriyorlar. Bunlar ilaç araştırması niteliğinde olmayan, ancak aslında bir kısmı bilimsel, etik vb. yönlerden çok daha büyük sorunlar taşıyan çalışmalardır. Biz bunları da değerlendiriyoruz. Buna karşılık aslında ilaç çalışmalarının sayıları çok da yüksek değildir. Gerçekten araştırıcı inisiyatifi ile başlatılmış olan çalışmaları değerlendirecek olan kurulların bağımsız olması gerekir.

Hasan Yazıcı: Diyelim ki Behçet hastalığında bir aspirin çalışması yapmaya karar verdik. Bunu topluma anlatmanın hiçbir sakıncası yok. Böyle bir çalışmaya gerek var mı? Yüzde yüz var! Çocukların kanser çalışmalarının hiçbiri ilaç firmalarından destek almıyor. Ama ilaç çalışması deyince toplumun aklına hemen başka şeyler geliyor. Niye? Bizim yüzümüzden geliyor. Biz kuralları böyle yapar isek gelir.

Cem Özesen: Klinik pratikte hekimlere hiçbir kısıtlama konmazken klinik araştırmalarda bu kadar kısıtlanma konmasının ve araştırmaların sadece ilaç firmalarının sponsorluğu ve onların yazdığı protokollerle yapılmasının tehlikeli bir şey olduğunu düşünüyorum. Çünkü ben hastayı tedavi ederken bir hastaya bir ilacı veriyorum ve "niye verdin?" diye sorulduğunda "faydalı olacağını düşündüğüm için verdim," dediğim anda iş bitiyor. Şu anda bu uygulamada bir hastaya bir ilacı vermek serbesttir; hiçbir kısıtlama, sigorta etme filan gerekmiyor.

Bir Katılımcı: Hayır alabiliyorsun. Çok az ilaçta endikasyonuna bakılıyor. Yani hastayı tedavi ederken hiçbir kısıtlamamız yok; yani "ben bu hastaya iyi geleceğini düşündüm," dediğiniz anda, "daha önce başka hastalarıma verdim," ya da "hocamdan duydum, o bir hastasına vermiş, iyi gelmiş," dediğim anda hukuki bir sorun yok. Ama aynı şeyi bir klinik araştırmada yapacağım dediğim zaman o kadar çok yükümlülüğü var ki! Yani bu biraz paradoksal. Ben klinik araştırma için hastaya iyi gelir düşüncesiyle her şeyi verebiliyorum, araştırma yaparken çok kısıtlanıyorum ve ancak endüstri isterse, para verirse, protokolünü yazarsa yapabiliyorum. Yani bu kuralların biraz düzeltilmesi gerekir.

Haner Direskeneli: O noktada etik komite veya Sağlık Bakanlığı bunun ne kadar iyi niyetle yapıldığını bilemez. Örneğin, yeni bir immünosüpresif ilacı hastalarda denemek istiyorsun; çok bilimsel nedenle de olabilir bu, ancak o immünsüpresifin eskisinden 100 kat pahalı olmasıyla da olabilir. Sıklıkla biliyoruz ki, özellikle bu gözlemsel nitelikteki açık çalışmalar Türkiye'de bu tür motivasyonlarla yapılıyor. Bir de şu tehlikesi var: Bunların bilimsel yönü zayıfsa literatürü de etkiliyorlar; yani literatürde bu tür gözlemlerden elde edilen birçok sonuç bulunuyor. Şimdi burada senin dürüstlüğün ne kadar değerlendirilebilir? Kaçınılmaz olarak bir ölçüde regülasyon olmak zorunda; yani senin hastaya tedavi amaçlı vermenle aynı şey değil.

Bir Katılımcı: Klinik araştırmalar ya da randomize klinik çalışmaların klinik pratiği iyi yansıttığına katılmıyorum. Çünkü klinik çalışmalara bakarsan metotreksatı kullanmamak lazım, klinik araştırmalarına bakarsan metotreksatın etkisi hem az, hem de yan etkisi çok fazla. Romatoid artritte sık sık karaciğer fonksiyonlarını kontrol etmen lazım, ama biz klinik pratikte kullanıyoruz. İlaç endüstrisi para yatırdığı ilaçların çalışmalarını öyle tasarlıyor ki etki en iyi, yan etki en düşük çıkıyor. Bir örnek vereyim: metotreksat etkisiz deniyor, bunu herkes söylüyor. Halbuki metotreksatın etkisiz olduğunu söyleyemeyiz çünkü çalışma ancak uygun şekilde çalışılmamış diyebiliriz.

Bir Katılımcı: Ben Dokuz Eylül Üniversitesi'nde İlaç Araştırmaları Etik Kurulu'ndayım. Etik kurulun sorunlarından biri araştırıcının tanımı. Beş yıllık uzman olmasının gerekmesi yüzünden gerçekten bir araştırıcının üstünde 15 tane çalışma olabiliyor ve tabii ilaç firmalarının düzenlemiş olduğu araştırmalarda çalışmacı olarak çalışıyor aslında ve bazen çalışma hakkında çok fazla bilgisi de olmuyor. Örneğin etik kurul bir şeyler soruyor ama bilgisi yok. Gerçekten bu yeni düzenleme buna bir çözüm getirebilir.

Fatoş Önen: Bizim etik kurulumuzda bir hukukçu var ve devamlı katılıyor toplantılara; aslında son derece faydalı da oluyor gerçekten, özellikle sigorta konusunda. Üniversitenin hukuk fakültesinden bir görevli, bir öğretim üyesi, her toplantıya geliyor ve oldukça da faydalı oluyor. Özellikle bu sigorta poliçelerinin değerlendirilmesinde, hastaların sigortalanması konusunda oldukça yararlı oluyor. Bunun dışında tabii etik kurulun en zorlandığı konulardan biri plasebo kontrollü çalışmalar. Bizim etik kurulda şöyle bir eğilim var: plasebo kontrollü çalışmaları kabul etmeyelim. Şöyle bir düşünce var: büyük ilaç firmaları bu tür ilaç çalışmalarını artık geri kalmış ülkelerde yapıyorlar, onun için bizi kullanıyorlar; bu önemli bir sorun.

Hasan Yazıcı: İlaç şirketi bir hipertansiyon ilacı çıkartıyor diyelim ki; onu önce plaseboyla sınamak istiyor. Etrafta diüretik varken, başka ilaçlar varken neden plaseboyla sınamak istiyor? Diüretikle sınamak istemiyor, çünkü 5 sene sonra diüretikle sınandığı zaman diüretikten daha iyi olmadığının görülmesi mümkün. Biz plasebo çalışmasının mantığını ve etiğini bilmiyoruz. Sevgili arkadaşlar bu çok hassas bir konu ama önemli bir konu. Yapılacak şey endüstriyle birlikte çalışmak, endüstriye dönüp iki tarafın birbirleri için yararını ortaya koyup birlikte çalışmak lazım. Bunun pazarlığına oturup karşılıklı konuşarak yapmak lazım. O pazarlık sonunda hem bizim -belki başta bizim- hem de endüstrinin etiği düzelecek gibi geliyor bana ve toplum bundan yararlanacak.

Haner Direskeneli: Bizimle endüstrinin çıkarlarının örtüşebileceğine inanıyor musunuz? Bilim olarak doğruyu aramak anlamında burada bunu tartışıyoruz; ordaysa ticari amaçları, ekonomik çıkarları olan bir kuruluş var. Bu işbirliği nereye kadar olabilir?

Hasan Yazıcı: Olabilir sevgili arkadaşlar. Şimdi listelere bakalım; şu anda kullandığımız

ilaçların çok büyük kısmı efektif ilaçlar, endüstrinin üniversiteyle beraber işin içine girip kullandığı ilaçlar. Endüstrinin bize çok yardımı var. Endüstrinin hakikaten büyük katkısı var ama karşılıklı pazarlığımızı yapıp birlikte çalışmamız lazım. Yegane yolu bu ama Avrupa pazarlığını yapmıyor, Amerika'da pazarlığı bazı yerler yapıyorlar. Şu andaki bütün etik kurallar işte hekimleri buraya götürdüler, yemek verdiler vs. ile ilgili. Halbuki başka sorunlar bence çok daha önemli.

KAYNAKLAR:

1. Suntharalingam G, Perry MR, Ward S, Brett SJ, Castello-Cortes A, Brunner MD, Panoskaltsis N. Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. N Engl J Med 2006; 355: 1018-28.
2. World medical association declaration of Helsinki. Ethical principles for Medical research involving human subjects. www.wma.net/e/policy/pdf/17c.pdf
3. Temple R, Ellenberg SS. Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 1: ethical and scientific issues. Ann Intern Med 2000; 133: 455-63.
4. Ellenberg SS, Temple R. Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 2: practical issues and specific cases. Ann Intern Med 2000; 133: 464-70.5.
5. Operational Guidelines for Ethics Committees That Review Biomedical Research, WHO, 2000. www.who.int/tdr/publications/publications/pdf/ethics.pdf

HAYAL ÜRÜNÜ BİR ÇALIŞMA: GUT TEDAVİSİNDE E VİTAMİNİ
Plasebo Kontrollü bir Çalışma

Yanlış Bulma Egzersizi

Özet

Vitamin E'nin gut tedavisinde faydalı olabileceğine dair gözlemler mevcuttur. Biz de, 6 aylık, çift kör, plasebo kontrollü bir çalışma ile 200 mg/gün ve 100 mg/gün olmak üzere iki ayrı dozu araştırdık. Vitamin E'nin her iki dozu da, ürik asit kan düzeylerini düşürme konusunda, plaseboya üstün değildi.

Giriş

Birkaç yeni çalışma, vitamin E'nin serum ürik asit seviyelerini düşürebileceği yönünde ipuçları vermiştir. Bunlar ya tek vaka bildirimleri veya ufak hasta gruplarında yapılmış, kontrolsüz, açık çalışmalardır. Vitamin E'nin pürin sentezini etkileyebileceğine dair gözlem oldukça eskidir. 1930'larda, Bocheim, vitamin E'nin ksantin oksidazı bloke ettiğini ileri sürmüştür. Sonrasında, başkaları, asetil koenzim A'nın çok etkili bir uyarıcısı olduğunu ve böylelikle, pürin sentezini bloke ettiğini öne sürmüşlerdir.

Bu nedenle, vitamin E'nin, gut tedavisinde etkili olabileceği hipotezini, 6 aylık, çift kör, plasebo kontrollü bir çalışmayla araştırdık. Ayrıca, vitamini, etkili olduğu savunulan, iki ayrı dozda araştırdık.

Hastalar ve Metodlar

Hastalar, her ikisi de İstanbul'da bulunan, Cerrahpaşa Tıp Fakültesi Romatoloji Ana Bilim Dalı ve Marmara Üniversitesi Tıp Fakültesi Diyabet ve Metabolik Hastalıklar Ana Bilim Dalı polikliniklerine ayaktan tedavi için düzenli olarak gitmekteydi. Bu hastaların bir kısmı, ürik asit seviyelerini düşürmek amacıyla geçmişte vitamin E kullanmışlardı. Bu hastalardan vitamin E kullanımı 3 ayı geçmiş olanlar çalışmaya dahil edildi. Her iki gruptan da gut tanısı almış ilk klinik başvurularında ürik asit seviyeleri 8.0 mg/dl üzerinde olan 50 yaşın üzerindeki erkek hastalar çalışmaya alındı.

Hastalar, çalışmaya katıldıkları sırada kullandıkları ilaçları almaya devam ettiler. Çalışma boyunca gut ataklarını engellemek amacıyla, ibuprofen 1800 mg/gün 10 gün süreyle uygulandı.

Denekler üç gruba randomize edildi. Grup A (n=36): plasebo; Grup B (n=36): 100 mg/gün vitamin E ve Grup C (n=36): 200 mg/gün vitamin

E. Her üç ilaç birbirinin aynı tablet formundaydı ve hastalar tarafından kahvaltıdan 30 dakika önce, günde birkez, aç karnına alındılar.

Çalışma her iki hastanenin etik komitesi tarafından onaylandı. Her hasta, randomizasyon sonrasında, bilgilendirme ve onay formu imzaladı. Hastaların, her ay, serum ürik asit, hematokrit, serum kreatinin, kolesterol ve testosteron seviyeleri incelendi. İstatistik analizi için Student's t testi kullanıldı.

Sonuçlar

Tablo 1, hastaların çalışmaya alındıkları dönemdeki özelliklerini göstermektedir. Tüm grupların, yaş, ürik asit seviyeleri ve hastalık süreleri açısından benzer oldukları görülmektedir. Tablo 2 hastaların çalışma başında ve sonundaki ortalama ürik asit, kolesterol ve testosteron seviyelerini göstermektedir.

Her üç grupta da, ürik asit seviyelerinin çalışma sonunda anlamlı olarak düştüğü görülmektedir (Grup A için p = 0.008; Grup B için p = 0.007; ve Grup C için p = 0.005).

Serum kolesterol seviyelerinde anlamlı fark gözlenmemiştir. İlginç olan, Grup A'dan bir, Grup B'den 3, Grup C'den 6 hasta, uzun yıllardır mevcut olan çarpıntılarına, vitamin E'nin iyi geldiğini düşündüler. Oniki hasta değişik sebeplerle çalışmayı bırakmak zorunda kaldı. Sonuç olarak Grup A'dan 6, Grup B'den 4 ve Grup C'den 2 kişi çalışma tamamlanmadan bıraktı (NS). Böylece, ilacın etkinliği konusundaki inceleme 96 hasta ile yapıldı. Grup A'da bırakmak zorunda kalan 6 hastanın 3'ü başka şehirlere taşındı, diğer 3'ü hakkında bilgi edinilemedi. Grup B'de 2 kişi dispepsi nedeniyle çekildi, Grup C'de 1 hastanın tedaviyi bırakma nedeni baş ağrılarıydı.

Tartışma

Gut tedavisinde vitamin E kullanımına ait bu ilk, çift kör, randomize plasebo kontrollü çalışmada, vitamin E'nin plaseboya üstün olmadığı tespit edildi. Ürik asit düzeylerini bir miktar düşürmüş olmakla beraber, özellikle 200 mg/gün dozda, bu bulgu, plasebodan anlamlı derecede farklılık göstermedi. Ancak, bu durum çalışmada az sayıda hastanın incelenebilmiş olmasına bağlı olabilir—Tip II hata.

Sonuçlarımızı açıklayacak bir nokta, *Hawthorne bias*'ı olabilir. Serum ürik asit seviyesinin azalmasının altında yatan asıl neden, her muayenede hastaların diyetleri konusunda sorgulanmış olmaları olabilir.

Bu çalışmadaki diğer önemli gözlem, kan basıncıyla ilgiydi. Her muayenede kan basıncı rutin olarak ölçüldü. İlginç bir şekilde, Grup C' deki 12 hipertansif hastanın 10'unda kan basıncı normalleşti. Vitamin E'nin kan basıncı üzerine olumlu etkisi Grup B'de de görüldü. Grup A'daki hastalarda bu tür etkiler saptanmadı.

Özetle, çalışmamız vitamin E'nin gut tedavisinde faydalı olduğu hipotezini kanıtlayamadı. Ancak, yine de bir miktar etkinlik gözlendi; bu konuda daha fazla sayıda hastayla çalışmalar

Tablo 1. Çalışmanın başında, hastaların yaş, hastalık süreleri ve serum ürik asit seviyeleri (ortalama ± SS)

Gruplar	N	Yaş (yıl)	Hastalık süresi(yıl)	Ürik asit(mg/dL)
A (plasebo)	30	58 ± 4.2	10.4 ± 2.4	9.3 ± 1.2
B (vit E 100 mg)	32	59 ± 3.8	12.8 ± 3.2	9.4 ± 0.8
C (vit E 200 mg)	34	61 ± 5.1	9.6 ± 2.9	9.8 ± 1.0

Tablo 2. Çalışmanın başında ve sonunda ürik asit, kolesterol ve testosteron seviyeleri (ortalama ± SS)

	Ürik asit (mg/dL)		Kolesterol(mg/dL)		Testosteron(nmol/L)	
	Başlangıçta	Sonunda	Başlangıçta	Sonunda	Başlangıçta	Sonunda
Grup A	9.3 ± 1.2	7.2 ± 0.8	266.4 ± 38	262.3 ± 12	24.52 ± 3.75	25.83 ± 3.36
Grup B	9.4 ± 0.8	6.9 ±1.1	258.8 ± 13	270.6 ± 14	23.83 ± 7.04	26.78 ± 3.92
Grup C	9.8 ± 1.0	6.7 ± 0.9	270.1 ± 18	280.4 ± 9	28.42 ± 5.25	30.10 ± 3.01

yapılmalıdır. Vitamin E zararsız bir ajandır ve bu çalışmada tesadüfen saptanan antihipertansif etki gibi ilave yararları düşünüldüğünde, bu tür çalışmalar kesinlikle gereklidir. Biz de, kurumumuzda bu doğrultuda bir çalışmayı başlatmak üzereyiz.

KAYNAKLAR

http://www.journalimportantresults.com/manuscript66

ÖRNEK İLAÇ ÇALIŞMASI SUNUMU: KATILIMCILARLA BİRLİKTE DEĞERLENDİRME

Vedat Hamuryudan: Hepinizin okuduğunu varsayarak üzerinden geçeceğiz, belli majör hataları bulmaya çalışacağız. Sizin dosyalarınız içinde bir kırmızı bir sarı kart var. Onları kullanalım, ayrıca ben bazen sorular soracağım. Kırmızı hayır sarı evet anlamına gelecek. Arkadaşlar ilk soruyu soruyorum: Farz edelim bu yazı bir dergiye geldi, siz de editörüsünüz. Bu yazıyı kabul ediyor musunuz? Bu çalışmayı plasebo kontrollü bir kör çalışma olarak tanımlasam kabul eder misiniz? Ne olarak tanımlıyorsunuz?

Bir Katılımcı: Bir yerinde hastaların randomize edildiği yazılmış.

Vedat Hamuryudan: Yazılmış. Bu yeterli mi?

Bir Katılımcı: Randomizasyonun nasıl yapıldığı yazılmamış.

Vedat Hamuryudan: Yani bana güvenmiyor musun? Genel olarak makalelerde ben bunu randomize yaptım diye açıklamak yeterli değil mi? Bunu randomize yaptım demek yetmiyor, randomizasyon metodunu da açıklamak gerekiyor. Peki hasta oluru alınmış mı?

Bir Katılımcı: Alınmış ama randomizasyondan sonra alınmış.

Vedat Hamuryudan: Ona niye itiraz ediyorsun?

Bir Katılımcı: Öncesinde alınması gerekiyor.

Vedat Hamuryudan: Öncesinde alınması, yani randomizasyondan önce çalışmanın hastaya açıklanması ve onayının alınması lazım.

Bir Katılımcı: Çünkü randomizasyon yaptığınızda hastayı zaten çalışmaya sokmuş oluyorsunuz. Sokmadan önce onayını almamız gerekiyor.

Vedat Hamuryudan: Bu çalışmada etik bakımdan sıkıntı var mı? Plasebo verilmiş, hastalar 6 ay süreyle plasebo kullanılmış. Bir de etkin tedavi var. O etkin tedaviyi 6 ay süreyle vermemenin uzun dönem prognozu değiştirdiğini veya hastalığın seyrini değiştirdiğini söyleyen bir veri var mı? Bildiğimiz kadarıyla gutta 6 ay kolşisin kullanmazsan, bu hastalar 10 yıl sonra çok kötü olurlar diye bir veri yok. Peki, atağı olan hastayı plaseboya vermek bir sıkıntı yaratır mı? Burada o bakımdan etik bir sorun var mı?

Bir Katılımcı: Plasebo alan grup daha fazla atağa girebilir, o açıdan sıkıntısı olabilir.

Vedat Hamuryudan: Ataklarda allopurinolü kesiyor musunuz?

Bir Katılımcı: Hayır,

Vedat Hamuryudan: Kesmiyorsunuz. Peki, hasta grubunun tanımlanması açısından bir eleştiriniz var mı? Allopurinol almayan grupta yani plasebo alan grupta atak sıklığı artmış olabilir, hasta bu nedenle atağa girmiş olabilir ama sonuçta atak tedavisini almış oluyor değil mi? Çalışmaya sadece erkek hasta almış; bu kabul edilebilir bir şey mi? Gut hastalığı deyince aklıma nasıl olsa erkek geliyor, belki ben bunu kabul edebilirim ama burada en azından bir açıklama beklerdim; yani niye böyle bir cinsiyet seçimi yaptı? Sırf kadınlarda da yapabilirdi bunu.

Bir Katılımcı: Tüm grubun ortalama olarak verilen değeri aslında büyük oranda erkeklerin verilerini yansıtacak; zaten kadınların orada, değer üzerinde etkinliği olmayacaksa grubu tamamen homojen hale getirmek için bu grubu seçmiş olduklarını bildirmeleri yeterli olabilir diye düşünüyorum.

Banu Çakır: Cinsiyetin sonuçlar üzerinde etki yarattığını düşünüyorsanız, vitamin E kullanımıyla ürik asit oluşması arasındaki etki ya da atak gelmesi arasında etken ve sonuç ilişkisini cinsiyetin karıştırdığını düşünüyorsanız sadece bir cinsiyette yapabilirsiniz. Cinsiyet karıştırıcı faktör mü değil mi onu hiç değerlendirme şansınız yok ama baştan karıştırıcı faktör olduğunu kabul edip onun etkisini silmiş oluyorsunuz. Bir tanesi de etki değişimi olabilir. Dün hatırlarsanız dedik ki, önce etki modifikasyonu var mı diye bakarız, yoksa zaten karıştırıcı faktör mü diye bakıyoruz.

Eğer cinsiyetin alt gruplarında vitamin E'nin ürik asit depolanmasıyla ya da gutla ilgisinin farklı olduğuna inanıyorsanız, kadınlarda iki kat etkili de erkeklerde beş kat diye düşünüyorsanız tabakalı analiz yapıyoruz. Bu da bir tabakalama yöntemi; ben sadece erkeklerdekini bilmek istiyorum, kadınlar benim çalışmamı etkiler diye düşünmüş olabilirsiniz. Ama bunların hepsinin bir şekilde mutlaka yazılmış olması lazım çünkü aksi takdirde yanlı bir çalışma olduğunu düşünmemiz gerekir.

Vedat Hamuryudan: Burada hasta tanımına başka bir itirazınız var mı ilaç kullanma açısından? Hastaların bazıları daha önceden zaten denenmekte olan ilacı kullanmışlar. Bu hastaların bir kısmı daha önceden vitamin E'yi bu amaçla almış olabilir. En az 3 aydan beri almıyorsa, daha önceden kestiyse bu hastaları ben alıyorum. Bu sizce kabul edilebilir bir şey mi?

Bir Katılımcı: Vitamin E'ye nasıl cevap verdiklerini bilmiyoruz, bunlar cevap vermemiş kişilerse o zaman *under-matching* olmuş olacak.

Hasan Yazıcı: Kural olarak daha evvelden ilaçları kullanmış olsalar da, kullanmamış olsalar da sorun var. Eğer yalın araştırma yapmak istiyorsak kullanmamışlardan seçmek lazım. Kural olarak bilimsel araştırma yapmak istiyorsak daha önce ilaç kullanmamış hastaları almak lazım, çünkü önceden ilacı kullanmış hasta ilacı etkisinden veya yan etkisinden tanıyarak kör olmaktan çıkabilir.

Vedat Hamuryudan: Anti-TNF bir ajanla yapılan yeni bir çalışma var, 2–3 sene önce yayınlandı. Bu çalışmada metotreksat kullanılmamış, yani metotreksat naif olarak tanımlanan hastalar çalışmaya alınmış, ama parantez içinde daha önceden kullanmış ama 6 aydan beri kullanmayan hastaları da naif kabul etmişler. Çalışmanın sonuçlarında bu naif hastaların % 45'inin daha önceden metotreksat kullanmış olduğu belirtiliyor.

Bir Katılımcı: Burada hipotez gutun tedavisi ama sonuçlarda ürik asit seviyesi, kolesterol vs. alınmış.

Vedat Hamuryudan: Doğru! Yani outcome işin içine katılmış. Oysa bu yöntemde anlatılmamış. Neyi ölçtüğü tanımlanmamış. Bu çalışmanın hedefinin ne olduğu tanımlanmamış. Sonra güç analizi yok.

Banu Çakır: Ben bir şey soracağım: Ortalama kaç ilaç kullanıyorsunuz böyle bir gutlu hastada, hasta grubunda? Sizin hastalarınız kronik hastalar olduğu için çok ilaç, değil mi? Burada hastanın daha önceki ilaçlarının devam ettiğini söylüyor, aslında bu da çok önemli: o devam ettiği ilaçlar hangileri, onlardan herhangi birinin bir etkileşimi olabilir mi? Ama onun ötesinde, burada hiç yok ama, hastanın uyuncu açısından gerçekten süre arttıkça problem olacak; hastanın aldığı diğer ilaç sayısı ve cinsinin de belirtilmesinde yarar var. O da atlanmış.

Hasan Yazıcı: Burada *power* analizi hakkında yaygın bir günahtan bahsedeceğim. Son zamanlarda iyi dergilerde görüyorsunuz; "bu çalışma % 84 veya 86 güçle yapıldı," diye belirtiliyor. *Power*, daha evvelden hesaplanmışsa, % 80 veya % 90 diye hesaplanır; % 84, 86, 87 diye yapılmaz. Bunların çok büyük bir kısmının daha sonradan yapıldığı çok açık ve doğru söylenmiyor. Hiçbir araştırıcı % 84 *power* ile hesap yapmaz ama böyle yapmayı tercih ediyorlar.

Banu Çakır: Elinizde vaka sayısının az olduğu durumlarda bazen *power* analizi yapmadan da çalışmaya başlayabilirsiniz ya da kabaca yaparsınız. Örneğin; %90 *power* diye yola çıkarsınız ama ulaşamazsınız o sayıya ya da hastaların bir bölümü çalışmayı tamamlayamaz. O zaman bir reverse *power* analizini destekliyoruz. *Power* analizi, bir şey bulamadığınız zaman önemlidir, bulduğunuz zamandan ziyade bulamadığınız zaman önemlidir. Gruplar arasında burada da fark bulunamamış; bulunamadığı zaman benim editör olarak ilk soracağım şey şu olacaktır: "Bunun *power* analizi neydi?" O aşamaya kadar yapmadıysanız bile, o aşamada beni editör olarak inandırmanız gerekir. İkinci olarak, NNT diye bir kavram var: *Number Needed to be Tried*. Klinik araştırmanın her bacağında kaç kişi olması gerekiyor ki ben bu sayıya ulaşayım. Özellikle burada tüm gruplardan ciddi bir *drop out* var ve dengeli değil; dolayısıyla yine burada o drop out'lar nedeniyle hedeflenen sayıya ulaşılamadığı için de bir şey bulunamamış olabilir.

Vedat Hamuryudan: Yönteme başka itirazınız var mı?

Hasan Yazıcı: Seçme sırasında hastaların alınmasında konulan bir koşul var: Serum ürik asit düzeyinin 8 mg'dan yüksek olması. Eğer ben

100 tane gutlu hasta tedavi ediyorsam, ortalama 7 civarındadır, 7,5 civarındadır. Böyle olması bize bir sorun yaratıyor mu? Yaratıyor. Dün bahsettiğimiz *regression to the mean* olgusunu yaratıyor. Dokuz mg'ın üzerindekileri alsaydım, bir fark gösterebilmek için daha büyük grup çalışmam lazımdı. On alsaydım daha da çok hasta almam gerekecekti.

Banu Çakır: Ben doğrudan T testiyle ilgili bir şey söylemek istiyorum. Bu makalede yok ama, bizim Türkiye'de yaptığımız ve yurtdışı yayınlara gönderirken eleştiri aldığımız çok ciddi bir hatadan da bahsedeceğim. Özellikle Türk dergilerinde yayınlanan yazılar için söylüyorum; istatistik analizlerinde "SPSS Versiyon 11 paket programıyla yapılmıştır," gibi bir cümle oluyor. Bunun hiç anlamı yok.

Bir Katılımcı: Bazen editör istiyor, hangi paketin kullanıldığının belirtilmesini.

Vedat Hamuryudan: Bu tür bir çalışmada ilacın farmakokinetiğini de göstermek bakımından, mesela bireyler arası serum konsantrasyonlarını da işin içine sokmak acaba daha doğru olur mu?

Bir Katılımcı: Hastaların ortalama yaşı verilirken +/- 'den sonra 4.2 şeklinde küsuratlı verilmiş, sadece 4 yazmak yetmez mi?

Vedat Hamuryudan: Yuvarlamadan bahsediyorsun. En yakın tama yuvarlamak. 4.24 yerine 4.2 şeklinde verebilirsin.

Bir Katılımcı: Seçilen yaş grupları arasında istatistik bir anlamlılığın olmadığının belirtilmesi gerekiyor. Mesela 58–61 arasında böyle bir şey var mı, yok mu onu bilmiyoruz.

Vedat Hamuryudan: Evet. Bütün gruplar benzerdi deniyor ama tabloda da ayrıca belirtmek lazım.

Bir Katılımcı: Tablo 1'i vermeden önce gruplardan ayrılan hastaların ayrı, o aşağıda belirtilen açıklamaların başta olması lazım.

Vedat Hamuryudan: Yöntemde soru gelmesi amacıyla bilerek bir hata yaptım. Hastalık süresi deniliyor; biz bu hastaların hastalık sürelerinin nasıl hesaplandığını biliyor muyuz? Yöntemde tanımlanmış mı bu? Gutun ne olduğu tanımlanmış mı? Nasıl gut denmiş? Doktor demiş, hasta demiş, eczacı demiş, kriter uygulanmış, hasta grupları hakkında bir bilgi var mı? Yani bir atak geçirmiş, bir daha hiç geçirmemiş mi? Kronikleşmiş mi? Bunlar yok ve burada bir hastalık süresi karşımıza çıkıyor. Bu da belirsiz şeylerden bir tanesi. İkinci tabloya geçecek olursak, bu tablo bize neyi gösteriyor? Ürik asit, kolesterol ve testosteron değerlerini.

Banu Çakır: Aslında size çok basit gelebilir ama klinik ortamlarda genellikle sordukları için soruyorum, herkes standart sapma ya da standart hatanın ne zaman kullanılacağından huzurlu mu? Mesela burada standart sapma verilmiş, benzer şekilde bir başka çalışmada standart hatasıyla birlikte verebilir; ortalamayı sık olarak standart hatasıyla da görüyoruz. Standart hata, verdiğiniz ortalamanın standart sapmasıdır. Standart hata verdiğiniz ortalama değerin çevresindeki dağılımın ne kadar olduğunu gösteriyor. Onun için biz sık olarak standart sapmayı veriyoruz.

Vedat Hamuryudan: Tartışma bölümüne gelirsek, yorumlarınız nelerdir?

Bir Katılımcı: Tartışmada birden bire hipertansiyonu görüyoruz, ama sonuçlarda hiç bahsedilmemiş. Ayrıca nasıl seyretmiş onu da bilmiyoruz.

Hasan Yazıcı: Makalenin son cümlesine takıldım. "Biz, bu konuda yeni bir çalışmaya başlıyoruz," diyor. Buldukları sonuçla ilgili başlayacakları çalışmanın haberini veriyor. Bunun yapılmaması lazım, onun için siz koymayın öyle cümleler, yeni bir çalışma yapacaksanız yapın!

A

aday gen yaklaşımı 50

add-on, bkz. tedaviye ek tedavi

advers olay 65, 66, 67, 76

advers olay bildirimi 65

advers etki 65, 76

aktif kontrollü çalışma 78

alfa değeri 18, 20

analitik çalışmalar 37

asosiyasyon 51

assay sensitivity, bkz. deney duyarlılığı

aşağı olmama 60

aşırı eşleştirme 32, 33, 35

Avrupa GCP'si 73

B

bağlantı analizi 46, 52

bağlantı dengesizliği 51, 52, 53, 55

Bayes teoremi 22

Berkson paradoksu 33

Berkson yanlılığı 32, 33, 34

Berkson's bias, bkz. Berkson yanlılığı

bias, bkz. yanlılık

bilgi toplama 32

bilgilendirilmiş olur 75

biyolojik gradyan 40

C

case-cohort 39, 40

Cochrane sistematik derleme veritabanını 60

collaborative analysis 42

conditional data 22

confounding, bkz. karıştırıcı faktör

Consort Bildirgesi 62

coverage error, bkz. kapsayıcılık hatası

cross-over 46, 60

cross-sectional çalışmalar 37, 38

Ç

çalışma koordinatörü 74

D

değişken 12, 13, 16, 18, 28, 29, 31, 35, 47

değişkenlik 31, 47

deney duyarlılığı 60

deneysel çalışmalar 37

design-based approach 17

dış geçerlilik 32

disease duration, bkz. hastalık süresi

doğru 32, 33, 34

dominant 46, 56, 57

drop out 90

duyarlılık 15

düzenleyici polimorfizm 49

E

early escape, bkz. erken kaçış

ecology bias, bkz. ekoloji yanılgısı

effect modification, bkz. etki değişimi

effectiveness, bkz. etkinlik

efficacy, bkz. etki

efficiency, bkz. verim

E. I. Cantekin 32

ekoloji yanılgısı 32, 34, 41

ekolojik çalışmalar 37

equivalence, bkz. eşitlik

erken kaçış 79

error, bkz. hata

eşitlik 60

eşleştirme 29, 32, 33, 35

etki değerlendirme ölçütleri 59

etki değişimi 26, 28, 89

etkileşim 26, 27, 28, 47, 50, 56, 89

etkinlik 59, 66, 79, 88

evidence-based medicine toolbox 40

exposure 38, 39

external validity, bkz. dış geçerlilik

F

faz 1 78, 80, 82

faz 2 76, 80

faz 3 76, 80

faz 4 66, 76, 80, 82

fenotip 46, 47, 48, 49, 50, 51, 52

fonksiyonel polimorfizm 51

Framingham çalışmaları 38

funnel plot, bkz. huni grafiği

G

GCP, Good Clinical Practice 71, 72, 73, 75, 81

gecikme saptırması 43

geçerlilik 9, 10, 12, 14, 15, 16, 18, 29, 31, 32

geçiş dengesizliği testi 52

gen ekspresyonu 35

genelleştirilebilirlik 17

generalizability, bkz. genelleştirilebilirlik

genetik çalışmalar 38, 39, 45, 46, 49, 50, 55, 56, 82

genetik epidemiyoloji 43, 45

genomda genel tarama 51

genotipleme hataları 58

gerçek 11, 12

geri çekme 79

girişimsel olmayan çalışma 82

Good Clinical Research Practice 71

gönüllü formu 81

gözlemsel çalışmalar 37, 59, 66, 67, 82

güç 14, 19, 36, 43, 52, 55, 90

güven aralığı 9, 14, 23, 43, 60

güvenlik 83

H

Hardy-Weinberg dengesi 57

harm, bkz. zarar
hastalık başlangıç zamanı 32, 34
hastalık süresi 32, 34, 88, 91
hata 9, 10, 11, 14, 16, 18, 19, 20, 22, 23, 31, 32 33, 35, 36, 40, 57, 58, 59, 60, 88, 91
hatırlama yanlılığı 32, 38
havuzlama analizi 42
Helsinki Deklarasyonu 78
heterojenite 41, 42
hibrid tasarım 39
huni grafiği 43
I
ICH GCP 71, 72, 73, 75
IND çalışma kuralları 73
İ
iç çoğalma 81
iç geçerlilik 16
İKU (İyi Klinik Uygulamalar) Kılavuzu 73, 74
ilaç araştırma dönemleri 83
İlaç Araştırmaları Yönetmeliği 80
ilişki çalışmaları 51, 52, 53
inbreeding, bkz. iç çoğalma
information, bkz. bilgi toplama
insan genomu 45
intention-to-treat, bkz. tedaviye kalkışma
interaction, bkz. etkileşim
internal validity, bkz. iç geçerlilik
investigator's brochure 75
işbirlikli analiz 42
İyi Klinik Uygulamalar 71
K
kalıtımsallık analizi 50
kapsayıcılık hatası 16
karıştırıcı faktör 24, 25, 26, 27, 28, 29, 89, 90
karıştırıcılık 29, 32, 35
keskinlik 12, 18
kesme noktası 20
Klinik Araştırmalar Danışma Kurulu 81
klinik araştırmalar yönetmeliği 73
kohort 16, 37, 38, 39, 40
kohort çalışmaları 37, 38, 39
kohort etkisi 39
kompleks genetik hastalık 47
konkordans 50
kontrollü ilaç çalışmaları 32
kümülatif meta analiz 42, 43
L
lambda 50
layman 81
lead time, bkz. hastalık başlangıç zamanı
likelihood 22
lod skoru 46.
logarithm of odds ratio, bkz. lod skoru
lojistik regresyon 16, 23, 29
M
matched-cohort 40
matematiksel olabilirlik 22
mathematical likelihood, bkz. matematiksel olabilirlik
meta analiz 41, 42 43, 66, 78, 79
meta regresyon 42
mikrosatellit polimorfizmleri 51
model-based approach 17
monitoring 74
monogenik hastalıklar 46, 47
multiplikatif 57
N
nakledilmiş dizi 45
negatif prediktif değer 15
NNT, number needed to be tried 90
non-inferiority, bkz. aşağı olmama
non-interventional study, bkz. girişimsel olmayan çalışma
null hypothesis, bkz. yokluk hipotezi
Nürnberg Mahkemeleri 72
Nürnberg Yasası 72
O
odds, bkz. olası risk
odds *ratio*, bkz. tahmini rölatif risk
olası risk 24
olgu kontrol çalışmaları 33, 34, 37, 38, 39
onaylanmış etik kurul 81
ortalamaya regresyon 32, 34, 35
outcome 40, 90
overestimation 23
over-matching, bkz. aşırı eşleştirme
P
p değeri 9, 11, 18, 20, 22, 23, 26, 29, 40, 57, 67
penetrans 39, 46, 47, 50, 53
plasebo 29, 33, 34, 36, 60, 61, 73, 79 80, 84, 87, 88, 89
plasebo kontrollü çalışmalar 60, 78, 79, 80, 84, 88
plasebo kullanımı 60, 88
pooled analysis, bkz. havuzlama analiz
popülasyon 12, 17, 66
post-marketing dönem çalışmaları 82
posttranslasyonel komodifikasyon 49
power, bkz. güç
pozitif prediktif değer 15
precision, bkz. keskinlik
prevalans 13, 15, 16, 18, 50
prevalans çalışması 15
promotor bölge 49
prospektif 39, 40, 42, 43
prospektif meta analiz 43
publication bias, bkz. yayın saptırması

R
randomize olmayan veri 66
rasgele hata 14, 18
recall bias, bkz. hatırlama yanlılığı
regression to the mean, bkz. ortalamaya regresyon
reliability 31
reproducibility, bkz. tekrarlanabilirlik
rescue 61
resesif 46, 57
retrospektif 39, 40, 42
reverse causation 38, 40
reverse power analizi 90
risk 12, 13, 14, 23, 24, 26, 27, 28, 37, 38 39, 40, 50, 53, 56, 57, 59, 61, 81
roving alpha 22
rölatif risk 13, 14, 23, 24, 27
S
safety, bkz. güvenlik
sağlıklı gönüllü 77, 80
sampling error 16
seçicilik 15
seçme 32, 35, 91
seçme yanlılığı 32, 35
seeding 82
selection, bkz. seçme
sibling değeri 50
sistematik derleme 42, 66
sistematik hata 11, 14, 18
SNP, single nucleotide polymorphism, bkz. tek nükleotid polimorfizmleri
standart hata 91
standart sapma 91
subgrup analizi 61, 62
superiority, bkz. üstünlük
surrogate 61
T
T testi 88, 91
tabakalı analiz 29, 90
tahmini rölatif risk 13, 24
TDT, transmission dysequilibrium test, bkz. geçiş dengesizliği testi
tedaviye ek tedavi 79
tedaviye kalkışma 60
tek nükleotid polimorfizmleri 48, 49, 52
tekrarlanabilirlik 31, 35
time lag bias, bkz. geciktirme saptırması
tip 1 hata 11, 20
tip 2 hata 11, 19, 88
toplum prevalansı 50
transkripsiyon 49
truth, bkz. doğru
Tuskegee Sifilis Çalışması 72, 73
U
unconditional data 22
under-matching, bkz. yetersiz eşleştirme
under-matching 32, 90
Ü
üstünlük 60, 80
V
validity, bkz. geçerlilik
variability, bkz. değişkenlik
variable, bkz. değişken
verim 14
withdrawal, bkz. geri çekme
Y
yan etki 40, 60, 62, 65, 66, 67, 68, 76, 77, 78, 82, 84
yanlılık 11, 29, 31, 32, 35, 41, 42, 55
yardımcı araştırıcı 74, 75
yayın saptırması 43
yetersiz eşleştirme 32, 33
yokluk hipotezi 18, 19, 20, 22, 23
yuvalandırılmış vaka kontrolü 39
Z
zarar 39, 40, 65, 66, 67, 68, 78, 79, 80
zarar bildirimi 65, 66
zarar sıklığı 66