

ROMATOLOJİDE KLİNİK ARAŞTIRMALAR VE NESNEL TIP II

EDİTÖR
HASAN YAZICI

A. GÜL
B. ÇAKIR
E. LESAFFRE
S. PAY
T. PINCUS
T. ŞOKKA
V. HAMURYUDAN
Y. YAZICI



ROMATOLOJİDE KLİNİK
ARAŞTIRMALAR VE
NESNEL TIP II

© AresKare

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

Ağustos 2008, İstanbul

Baskı ve Cilt: Promat Basım Yayın San. ve Tic. A.Ş.

Sanayi (Örnek) Mahallesi 1590. Sokak No: 32

34510 Esenyurt Büyükçekmece İSTANBUL

Tel : +90 212 622 63 63 • Fax : +90 212 622 63 73

ISBN 978-9944-214-02-5

İÇİNDEKİLER

Romatolojide Klinik Araştırmalar ve Nesnel Tıp II

Önsöz	5
Romatooid Artrit Neoplazi: Tedavi ile ilişkili mi?	7
<i>Anders Ekblom</i>	
“p” Değeri	9
<i>Emmanuel Lesaffre</i>	
Üstünlük, Eşdeğer Etkinlik ve Daha Az Etkili Olmama Çalışmaları (Superiority, Equivalence, Non-inferiority)	15
<i>Emmanuel Lesaffre</i>	
Randomize Kontrollü Çalışmaların Doğru ve Yanlış Kullanımı	23
<i>Hasan Yazıcı</i>	
İleriye Dönük Veri Tabanları ve Kayıt Sistemleri – Neden?	27
<i>Theodore Pincus</i>	
İskandinav Veri Tabanları	59
<i>Tuulikki Sokka</i>	
Türkiye Anti-TNF Veri Tabanı	63
<i>Salih Pay</i>	
QUEST – RA: Global Bir Veri Tabanı	67
<i>Tuulikki Sokka</i>	
ABD’de Romatizmal Hastalıklar için Veri Tabanları	73
<i>Theodore Pincus</i>	
Hasta İzlemi ve Veri Toplanması’nın Günlük Pratiğe Alınması – ARMD ve BOARD Veri Tabanları	75
<i>Yusuf Yazıcı</i>	
İnteraktif Oturum I	81
<i>Ahmet Gül</i>	
İnteraktif Oturum II	93
<i>Vedat Hamuryudan, Banu Çakır</i>	
İndeks	104

1. The first part of the paper is devoted to a general discussion of the problem of the existence of solutions of the system of equations (1) for arbitrary values of the parameters α and β .

2. In the second part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

3. In the third part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

4. In the fourth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

5. In the fifth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

6. In the sixth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

7. In the seventh part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

8. In the eighth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

9. In the ninth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

10. In the tenth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

11. In the eleventh part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

12. In the twelfth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

13. In the thirteenth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

14. In the fourteenth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

15. In the fifteenth part we shall consider the case of the system of equations (1) for arbitrary values of the parameters α and β .

ÖNSÖZ

Hasan Yazıcı

05

“II. Bodrum Romatolojide Klinik Araştırmalar ve Nesnel Tıp Günleri” 13-15 Nisan 2007’de yapıldı. Elinizdeki kitapçık bu toplantının ürünü. I. Bodrum Günleri kitapçığında da altını çizdiğim gibi “Araştırma bir yana, günlük hekimlik uygulamasında da nesnellik özellikle gelişmekte olan ülke hekimi için yaşamsal.” Bu cümleye “özellikle”yi bu yıl ekledim. Bodrum toplantımıza yine özellikle genç romatologların gösterdiği ilginin devamını diliyorum.

Tebliğli, tebliğsiz tüm katılımcılara, Romatoloji Araştırma ve Eğitim Derneği’ne (RAED) ve toplantımızın bilimsel içeriğine hiçbir şekilde karışmamakta örnek bir titizlik gösteren sponsorumuz Schering-Plough’a çok teşekkür ederim.

Düzenleme Kurulu adına
Hasan Yazıcı

Konuşma Özeti*

Kontrollü klinik çalışmalarda kanser gibi yan etkilere bakmak aşağıdaki nedenlerden ötürü oldukça zordur:

Kontrollü klinik çalışmalara katılan deneklerin nitelikleri gerçek hayatta görülen hastalara kıyasla çok değişiktir. Örneğin birden fazla hastalığı olanlar veya çok yaşlı veya gençler böyle çalışmalara pek katılmazlar. Kontrollü klinik çalışmalara katılan hasta sayısı hemen her zaman çok kısıtlıdır ve bu nedenle de seyrek rastlanan yan etkiler kontrollü klinik çalışmalarda pek görülmez. Bu nedenle neoplazi ve romatoid artrit ilişkisine başka türlü, yani kontrollü klinik çalışmalar dışında, toplanmış verilerden bakmak daha uygun olur.

İsveç sağlık sisteminde devlet tarafından oluşturulmuş ve tüm toplumu kapsayan veri tabanları bu iş için çok uygundur. İsveç'te 1947'den beri ulusal bazda veri toplanmaktadır. Biz de RA neoplazi ilişkisini söz konusu veri tabanlarına girerek inceledik. Esas incelemeler a. ulusal kanser veri tabanı ve b. ulusal hastaneye yatan hastalardan oluşturulmuş "tani veri tabanı" yardımıyla yapıldı.

On yıl ve daha uzun süre içinde RA tanısıyla hastaneye yatanlar arasında neoplazi sıklığına baktığımızda hastaneye yatışın ilk yılında neoplazi sıklığında bir artış vardı. Ancak artış izleyen yıllarda azalmış olarak sabit kalıyor, 10 yıl ve üzerine çıktığımızda ise serbest toplumdaki sıklıktan farksız hale geliyordu. Bu olgunun yorumunu ise birden fazla hastalığı olanın daha artmış sıklıkta hastaneye yatıp incelenmesinin neden olduğu saptırma olarak açıkladık. İlginç olarak RA'da solid tümörler arasında bazı kanser türleri artmış bazıları ise azalmıştı. Örneğin RA ile akciğer kanseri arasında bir birliktelik vardı. Sigara içmenin gerek RA gerekse de akciğer kanseri patogeneziindeki rolünü iyi biliyoruz. RA'da meme kanserinde bir azalma vardı. Bu da belki kronik iltihabın bazı kanser türlerinde azalmaya neden olması ile açıklanabilirdi. Bunun yanında mide kanseri sıklığı da RA'da biraz azalmış, kolon kanseri sıklığı ise belirgin (% 30) olarak azalmıştı. Kolon kanseri sıklığındaki azalmada doğal olarak RA'lıların kullandığı aspirinin önemli rolü vardı.

ROMATOİD ARTRİTTE NEOPLAZİ: TEDAVİ İLE İLİŞKİLİ Mİ?

Anders Ekbom

Karolinska Enstitüsü,
Stockholm, İsveç

Çalışmalarımızda RA ile lenfoma arasındaki ilişkiye baktığımızda önemli olarak aşağıdaki 2 sonuca vardık:

1. Hastalık şiddetiyle hastalarda non-Hodgkin lenfoma görülme sıklığında çok yakın bir ilişki vardı.
2. Anti TNF ajanlar da böyle ağır hastalara veriliyordu. O nedenle RA'da artmış olan lenfoma sıklığı ilaca bağlı değildi.

***Hazırlayan:** Konuşma özeti Hasan Yazıcı tarafından, ses kayıtlarından hazırlanmıştır.

GİRİŞ

P değeri, medikal çalışmalarda çıkarsama yapmak için en yaygın kullanılan araçtır. Bununla birlikte, popülerliğine rağmen, P değerinin önemi yeterince anlaşılamamaktadır ve medikal literatürde sıkça hatalı kullanılmaktadır. P değeri ile ilgili temel bilgiler, Bingham ve arkadaşlarından (2007) alınan örneklerle birlikte değerlendirilecektir. Bingham ve arkadaşları (2007), 12 hafta süre ile diz ve kalça osteoartriti tedavisinde standart tedavi olan selektif COX-2 inhibitörü etorikoksib 300 mg qd (ET) ve selekoksib 200 mg qd (CE) ile plaseboyu karşılaştıran, üç kollu, çift kör, randomize iki klinik araştırma bildirmiştir (Bingham et al., 2007). Her bir çalışmanın iki bölümü vardır. İlk çalışmada, 559 hasta ET, CE veya plasebo (PL) kollarına randomize edilmiştir. İkinci çalışmada, 608 hasta üç ilaçtan birini almak üzere randomize edilmiştir: ilk bölümü (etkin maddeye karşılık plasebo tedavisinin karşılaştırılması) başarıyla tamamlayan hastalar, ET ve CE'yi doğrudan karşılaştıran 14 haftalık takip çalışması olan ikinci bölüme dahil edilmiştir. Burada, klasik üstünlük çalışmasına odaklanılmıştır. Eşdeğerlik (equivalence) ve *daha az etkili olmama (non-inferiority) çalışması* için testler, bu kitabın 15. sayfasında yer almaktadır bkz. Lesaffre 2008). Primer sonlanım noktaları: (1) WOMAC Ağrı Alt-skalası (WOMACPA'nın ilk 5 sorusunun ortalamasıdır ve herbir soru VAS ile değerlendirilir; değerin büyük olması şiddetli ağrıyı ifade eder). (2) WOMAC Fiziksel Fonksiyon Alt-skalası (WOMACPH'nin 8'den 24'e kadar olan sorularının ortalamasıdır ve herbir soru VAS ile değerlendirilir; değerin büyük olması ağır hastalığı ifade eder) ve (3) Hastanın global hastalık değerlendirmesidir (PGADS-VAS ile değerlendirilir). Bu sonlanım noktalarının analizi başlangıçtan 12. haftaya dek, zaman-ağırlıklı ortalama değişime dayanmaktadır. Ayrıntılı bilgi orijinal yayında bulunabilir.

P DEĞERİ VE ANLAMLILIK TESTLERİ

P değeri kavramını açıklamak için, bu çalışmada başlangıçtan tedavinin 12. haftasına kadar ET ve PL arasındaki WOMACPA değişimini görelim. Bingham ve arkadaşlarının (2007) çalışmasında Tablo 2a'da ET ve PL için başlangıçta ve tedavi sonrasında sırasıyla ortalama 67.4 (66.6) ve 39.6 (54.2) WOMACPA

P DEĞERİ

Emmanuel Lesaffre

*Leuven Katolik Üniversitesi,
Belçika
Erasmus Tıp Merkezi, Rotterdam,
Hollanda*

değeri bildirilmektedir. Bu başlangıçtan itibaren ET için 27.8 ve PL için 15.4 'lük bir ortamda değişimi ifade eder, ET ve PL arasındaki başlangıçtan itibaren ortalama değişiklikteki (DA) fark, ET yönünde -15.07'dir. Bildirilen P değeri 0.001'den küçüktür ve bununla ne demek istediğimiz aşağıda anlatılmaktadır.

ET ve PL tedavilerinin eşit, yani ortalamalardaki gerçek farkın sıfır ($\Delta = 0$) olduğunu varsayın. Eğer $\Delta = 0$ ise, DA'nın sıfır civarı olmasını bekleriz ama aynı zamanda çalışma tekrarlandığında (sıfır civarında) dalgalanacaktır. Bu durumda sorunun şu olması gerekir: ET'nin PL'den performans olarak farklı olduğu sonucuna yüksek kesinlikte varmak için, DA ne zaman yeteri kadar büyüktür? Bununla birlikte, P değerinin hesaplanması için farklı bir soru sorar ve mevcut çalışma için, eğer $\Delta = 0$ ise, DA = -15.07 (veya mutlak değerde daha büyük) gibi bir sonucun ne kadar nadir olduğunu sorarız. PL'nin ET kadar iyi olduğu önermesi, yokluk hipotezi (*null-hypothesis*) olarak adlandırılır ve $H_0: \Delta = 0$ anlamına gelmektedir. $P=0.05$ değeri, DA = -15.07'nin H_0 (eğer Ho doğruysa) altında, oluşabilecek % 5 uç değerlere ait olduğu anlamına gelir. $P=0.01$ ($P<0.001$) olduğu zaman, yine elde edilen DA H_0 (eğer Ho doğruysa) altında, oluşabilecek % 5 uç değerlere aittir ve sadece sonuçların % 1'inde DA=-15.07 veya daha büyük bir farka eşittir (% 0.1'den az). P değeri hesaplaması, olasılık tartışmalarına dayanır ve istatistiksel anlamlılık testin sonucudur. İstatistiksel testler, yukarıdaki gibi yokluk hipotezinin (*null-hypothesis*) doğrulanması için tasarlanır. Genelde H_0 , reddetmek istediğimiz bir hipoteze karşılık gelir. Örneğimizde, PL'nin ET kadar iyi olduğunu reddetmek ve alternatif olarak, $H_A: \Delta \neq 0$ 'ın doğru olduğu sonucuna varmak isteriz. Hangi DA değerinden veya eş değer olarak hangi P değerinden itibaren H_0 'ı reddetmek istediğimize karar vermemiz gerekir. Alışılmış olan bu karar için eşik değer olarak, istatistiksel olasılık düzeyi $\alpha = 0.05$ 'in alınmasıdır. Özellikle:

- $P < 0.05$ olduğu zaman, elde edilen sonuç, H_0 'ın doğru olduğuna inanmak için çok fazla uç değerdedir ve H_0 reddedilir. Diğer bir deyişle, $\Delta = 0$ 'a inanmayız ve yukarıdaki örneğimizde olduğu gibi, 0.05'te sonucu istatistiksel olarak anlamlı kabul ederiz.

- Diğer taraftan $P \geq 0.05$ olduğu zaman, sonucun kabul edilebilen uç değerde olmadığı ve $H_0: \Delta = 0$ olabileceği dolayısıyla H_0 'ın reddedilemeyeceği sonucuna varırız. Bu nedenle, gözlenen farkın şansa bağlı olduğu ve $\Delta = 0$ olduğunu göstermek zorunda olmadığı sonucuna varırız. Sonucun $\alpha = 0.05$ 'te anlamlı olmadığını söyleriz. ET'yi CE ile karşılaştırırken, WOMAC fiziksel fonksiyon alt-skalası için $P = 0.367$ değeri istatistiksel açıdan anlamlı olmayan bir sonuç elde edildiğine işaret eder. Bu nedenle, ET'nin CE'den farklı olduğuna inanamayız.

Yukarıdaki alternatif hipotez H_A , $H_A: \Delta \neq 0$ 'ın, ya $\Delta > 0$ ya da $\Delta < 0$ olduğuna işaret ettiğinden "iki yönlü" olarak isimlendirilir. Bu durumda kullanılan istatistiksel teste ait P değeri de "iki yönlü" olarak adlandırılır. Randomize kontrollü bir çalışmada (RCT), çok sıklıkla testler iki yönlüdür, ve bu durum "bir tedavinin diğerinden daha kötü veya iyi olup olmadığını" tespit etmeye izin verir. Bazen alternatif hipotez (ve istatistiksel test, anlamlılık seviyesi, P değeri) "tek yönlü"dür; örneğin, $H_A: \Delta < 0$ veya $H_A: \Delta > 0$. Bununla birlikte, ruhsatlandırma kapsamında tek yönlü anlamlılık testi pek istenmez çünkü, bu yaklaşım bir tedavinin değeri üzerine daha etkin olduğunu kanıtlamayı amaçlarken, güvenlilik konularını tespit etmeye ortam sağlamaz.

Yanıt tipi, karşılaştırmanın amacı, deneysel ortam, sorulan sorunun cinsi vs. ile değişen çok çeşitli istatistiksel testler vardır. Mevcut istatistiksel testlerin özetini vermek, bu konuşmanın amaçlarının ötesine geçmektedir. Yine de, bazı iyi bilinen testler arasında (a) iki bağımsız (bağımlı) grubun ortamlamalarının karşılaştırması için ve verilerin her bir grupta eşit varyans ile normal dağılıma sahip olduğunun varsayılması; eşlenmemiş (eşli) t-testi (b) iki bağımsız (bağımlı) grubun iki oranın karşılaştırması için ki-kare testi (McNemar testi); (c) ikiden fazla bağımsız ortalamaların karşılaştırması için varyans analiz testleri; (d) Wilcoxon ve Kuruskal-Wallis testleri gibi t testinin aksine verilerin nasıl dağıldığına dayanmayan, nonparametrik testler sayılabilir.

$\alpha = 0.05$ şeklinde sonuç veren bir test aslında, $P < 0.05$ durumunda H_0 'ı reddettirir. H_0 'ın gerçekten doğru olduğu durumda bu reddediş yanlıştır. Böyle bir durumda, "Tıp I

hata” yapılmıştır. Diğer taraftan, “Tip II hata”, iki tedavinin gerçekte farklı bir etkisi olduğu durumda çalışmada H_0 ($P \geq 0.05$)’in reddedilemediğinde ortaya çıkar. Bir Tip I hata yapma olasılığı, P (Tip I hata), yanlış pozitif oran olarak adlandırılır. Tam olarak, anlamlılık seviyesi olan α ’ya eşittir ve bu nedenle, bir istatistiksel test (ancak sadece biri, aşağıdaki Çok Sayıda Test Problemi bölümüne bakınız) uygulandığında doğru sonuç verir. β ile gösterilen bir “Tip II hata” yapma olasılığı, otomatik olarak kontrol edilemese de, bu hatanın azaltılması için aşağıda göreceğimiz gibi örneklem büyüklüğünü ayarlamak gerekir.

P-DEĞERİNİN YANLIŞ KULLANIMI

1. Çok sayıda anlamlılık testi yapmak

H_0 doğru ise, istatistiksel anlamlılık seviyesini 0.05 olarak kullanan bir istatistiksel test yokluk hipotezini (*null-hypothesis*) 0.05’e tam olarak eşit olasılıkla reddedecektir. Böylelikle, gerçekte doğru olmasına rağmen, H_0 ’ın yanlış olduğunu iddia etmek % 5 risk taşır. Yüz çalışma yürütüldüğünü varsayın; bu durumda ortalama 5 çalışma için H_0 altında iki tedavinin farklı etkiye sahip olduğu şeklinde, yanlış sonuç ortaya çıkacaktır. Aynı zamanda, en az 1 çalışmadaki yanlış sonuç olasılığı, pratikte 1’e eşit olacaktır. Yanlış pozitif çalışma bulguları etkisiz ilaçların onaylanmasına yol açabileceğinden, ruhsatlandırma otoriteleri (FDA, EMEA) toplam yanlış pozitif oranının (en çok 0.05) sıkı sıkıya kontrol edilmesini mecbur kılar. Bu nedenle eğer çok sayıda test uygulanıyorsa, böyle bir yanlış sonucun ortaya çıkma olasılığını kontrol etmek için bir yöntem ihtiyacı vardır. Böyle bir düzeltme uygulanmazsa, çok sayıda test problemi ortaya çıkar. Çok sayıda anlamlılık testi yapmak aşağıdaki durumlarda sorun olur: a) İki tedavi yöntemi birden fazla ana sonuç (primary endpoint) açısından karşılaştırılıyordur. Birden fazla ana sonuçta istatistik anlamlılık bulunmuştur. Söz konusu osteoartrit örneğinde üç ana sonuç (WOMACPA, WOMACPH ve PGADS) vardır. b) İki den fazla tedavi yöntemi karşılaştırılıyordur. Osteoartrit çalışmasında üç tedavi yöntemi (ET, CE ve PL) vardır. c) Yaş, cins gibi çeşitli alt gruplarda alınan sonuçların karşılaştırıldığı durumlar. d) İki tedavi yönteminin etkilerini çalışma boyunca değişen zaman dilimlerinin sonunda kıyaslamak, e) Aynı sonuç için birden

fazla istatistik yapıp bunlardan sadece birinin anlamlı çıkması.

Çok sayıda test probleminin olası çözümleri şöyledir: 1) Birinci ama açıkça belli olan çözüm, sadece bir tek sonlanım noktası seçmektir ve dolayısıyla risk otomatik olarak 0.05’tir. 2) Eğer birden fazla sonlanım noktası düşünülmesi gerekirse, çok sayıda testten sakınmak için klinik veya istatistiksel gerekçelere dayanarak birleşik sonlanım noktası oluşturulabilir. Örneğin, akut MI hastalarının tedavi edileceği bir çalışmada, üç tane ikili sonlanım noktası yapılabilir: Kardiyak mortalite (E/H), infarktüsün tekrarlanması (E/H) ve tekrar hastaneye yatış (E/H). Böylelikle, en azından üç olaydan bir tanesi gerçekleşirse 1’e eşit ve hiç olmadıysa, sıfır olarak birleşik ikili sonlanım noktası yapılandırılabilir. Üç sonlanım noktasının birleştirilmesi, ağırlıkları bir istatistiksel prosedürle tayin edilen üç tane ikili sonlanım noktasının ortalama ağırlıklarının hesaplanması yoluyla sağlanabilir. 3) Çok sayıda testle uğraşmanın üçüncü yolu, onu düzeltmektir. Bu şöyle açıklanabilir: her test için ayrı olarak $\alpha=0.05$ olduğunu varsayın. En azından üç P değerinden bir tanesi 0.05’ten küçük olduğu zaman iki tedavinin istatistiksel olarak anlamlı derecede farklı olduğunu iddia edersek, iki tedavinin farklı etkinliği olduğunu yanlış olarak iddia etme riski yaklaşık olarak tek başına risklerin toplamına eşittir ve bundan dolayı da yaklaşık olarak 0.15’e (gerçekte 0.15’ten de küçük) eşittir. Toplam hata oranını kontrol etmek için (aile boyu diye de adlandırılır), bir seçenek; her test için anlamlılık seviyesini 0.05/3’e azaltmak ve bu vesileyle toplam hata oranını en çok 0.05’e azaltmaktır. Buna “Bonferroni ayarlaması” denir. Bu ayarlama k-testlerine uygulandığı zaman, iki tedavi arasında 0.05’te toplam anlamlı fark iddia edebilmek için her tek başına P değeri, α/k ’dan küçük olmalıdır. Bingham ve arkadaşlarının (2007) çalışmasında Bonferroni ayarlamasından sonra da önemli sonuçlar anlamlı kalmıştır.

Bonferroni ayarlamasından daha doğru yaklaşımlar, bazılarının kapalı test prosedürleri olarak adlandırdığı Simes, Holm, Hochberg ve Hommel’inkilerdir, (Moyé (2003)’e bakınız). Zaman içinde tekrarlayan durumlarda ise, örneğin ara analizler uygulanırken, toplam

Tip I hatanın kontrol edilmesi için buna tahsis edilmiş prosedürlerle çalışılmıştır. En popüler iki prosedür, O'Brien-Fleming ve Pocock'unkilerdir (Proschan, Lan ve Wittes (2006)'ya bakınız).

2. İstatistiksel açıdan anlamlı olmayan sonucun açıklanması ve örneklem büyüklüğünün hesaplanması

Çalışmalarda istatistiksel açıdan anlamlı olmayan sonuç durumunda ($P \geq 0.05$), yokluk hipotezi (*null-hypothesis*) reddedilmez. Böyle bir durumda, çalışma "negatif" diye adlandırılır. Böyle anlamlı olmayan bir sonuca rağmen, bazı araştırmacılar verilerde anlamlı sonuca doğru bir eğilim olduğunu idda ederek bir şekilde daha büyük örneklem ölçüsü ile çalışmanın kesinlikle pozitif olacağı yönündeki yanlış izlenime neden olurlar. Dahası sıklıkla bazıları anlamlı olmayan sonucu "iki tedavi arasında bir fark olmadığı"nın kanıtı olarak yorumlar. Şöyle bir deyiş vardır: "Kanıtın bulunamaması, olmadığının kanıtı değildir". Diğer bir deyişle, "bir fark gösteremeyiz" demek "ortada gerçekten de fark yok" demek değildir. Aslında, kimse istatistiksel araçlarla iki tedavi arasında eşit etki olduğunu gösteremez. Bu nedenle, yukarıda anlamlı olmayan bir sonuç durumunda "Ho kabul edildi" diye beyan etmekten kaçındık, bunun yerine "Ho'ı reddedemeyiz" diye belirttik.

Negatif çalışmalar sıklıkla örneklem sayısı çok düşük olduğu için ortaya çıkar ve dolayısıyla, tip II hata yapılma olasılığını çok artırır. Bu nedenle, çalışma planlayan kişi küçük bir β veya eşdeğer olarak "çalışmanın gücü" denilen büyük $1-\beta$ 'ya sahip olduğundan emin olmalıdır. Verilmiş bir örneklem büyüklüğü için güç, (a) primer sonlanım noktasına; (b) istatistiksel teste; (c) anlamlılık seviyesine ve en önemlisi de (d) iki tedavi arasındaki varsayılan klinik ile uyumlu farkı olan, Δ_5 'ye ("S" (Superiority) Lesaffre (2008) yayımında görüldüğü üzere iki tedavi arasında üstünlüğü ifade eder) dayanır. Güç (*power*), ayrıca, iki ortalamanın karşılaştırılması durumunda toplam standart sapmaya veya iki oran karşılaştırılırken kontrol hızı gibi diğer faktörlere de dayanır. Pratikte, gücün en az 0.80'le sonuçlanacağı örneklem büyüklüğü seçilir. Yüzde 80 ihtimalle, diğer değerlerin doğru olarak seçildiği varsayılırsa, verilmiş bir Δ_5 için yokluk hipotezinin

(*null-hypothesis*) reddedilebileceğini ifade eder. Örneklem büyüklüğünün pratikteki hesaplanması, oldukça tekniktir ve en iyi, bir yazılım paketi yardımı ile yapılabilir.

Bingham ve arkadaşlarının (2007) çalışmasında örneklem büyüklüğüne yönelik şu ifade vardır: "*Etorikoksib ve selekoksib gruplarının her birinde 200 hasta ve plasebo grubunda da 100 hasta ile yapılan çalışma, etkin maddeler arasında daha az etkili olmama ve etkin maddelerin plasebo üzerine üstünlükleri şeklindeki primer hipotezi sağlamak üzere, toplamda en az % 87 güç sağlamıştır*". Örneğin, iki ilacın (ET ve CE) plaseboya karşı kıyaslanması için Δ_5 'nin tanımlanması gibi, hesaplamaların tekrar edilmesi için yeterince detay vermediklerinde, örneklem büyüklüğü tanımları yetersiz kalmıştır.

3. İstatistiksel anlam ile klinik ilişki

Bingham ve arkadaşlarının (2007) çalışmasında Tablo 1, iki çalışmanın başlangıçtaki demografik verilerini gösterir ve her kol için açıklayıcı istatistik bilgiler vererek, üç kol arasındaki dengeyi değerlendirir. Dengesizliği kontrol etmek için istatistiksel bir test uygulanmamıştır ve bu doğru bir yaklaşımdır. Gerçekten de başlangıçta hiçbir hasta tedavi görmemiştir. O zaman diliminde hastalar, hiç tedavi görmeden sadece tedaviler için randomize edildiklerinden, tedavi etkisi olmaksızın yokluk hipotezinin (*null-hypothesis*) uygulandığından emin olabiliriz. Burada, kaç tane istatistiksel testin 0.05'te anlamlı olacağı tahmin edilebilirdi, örneğin 100 bağımsız testte ortalama 5. Açıkçası, anlamlı farklar beklenecektir ve bu nedenle de, başlangıçtaki dengeyi doğrulamak için anlamlılık testleri yarırsızdır.

İstatistiksel olarak anlamlı bir sonucun klinikle uyumlu olması gerekmez. Şöyle bir kurgusal örnek ele alalım: Akut MI hastalarını tedavi etmek üzere iki ilaç (A ve B) karşılaştırılıyor ve incelenen sonuç 1 yıllık mortalite. 400 hastanın iki katı ile bir çalışma başlatılıyor. Sonuçlar A için % 2 ve B için % 10 mortalite şeklinde ve ki-kare testi ($P < 0.01$) oldukça anlamlı bir sonuç veriyor. Bununla birlikte, aynı düzenekte 100,000 hastanın iki katının tedavi edildiğini ve sonuçların A: % 0.002 ve B: % 0.0010 olduğunu

varsayın. Yine ki-kare testi ile oldukça anlamlı bir P değeri elde edilecektir ama A'ya karşılık B'nin net klinik yararı şimdi önceki çalışmaya göre 100 kat daha düşüktür. Gerçekte her $\Delta \neq 0$ için, H_0 'ın yüksek olasılıkla reddedildiği bir örneklem büyüklüğü olduğu gösterilebilir.

(% 95) GÜVEN ARALIĞI

P değeri bir yandan randomize kontrollü çalışma sonuçlarını hala kritik olarak yorumlayan en popüler araç olmaya devam ederken, kullanımında ve yorumlanmasındaki zorluklar uluslararası literatürde giderek anlaşılmağa başlamıştır. Gerçekten de günümüzde *The Lancet* ve *New England Journal of Medicine* gibi kaliteli yayınlar P değeri yerine % 95 güven aralığını bildirmeyi tercih ediyorlar.

Örneğimizde gözlenen ortalama fark (DA =) -15.07, ortalamalardaki gerçek fark olan Δ 'nın bir tahminidir. Ama tabii ki, DA sadece bir tahmindir ve gerçek fark konusunda biz hala emin değilizdir. % 95 güven aralığı (GA) ise belirsizliğin ifadesidir. Daha spesifik olarak, % 95 GA, 0.95 olasılıkla gerçek fark Δ 'yı içeren aralıktır. Bu nedenle bu aralık dar olduğunda Δ hakkında iyi bir fikrimiz olur. WOMACPA için Bingham ve arkadaşları (2007) aşağıdaki farklar için % 95 GA bildirmişlerdir: (a) CE – ET: [-7.02, 0.77]; (b) ET – PL: [-19.72, -10.41] ve (c) CT – PL: [-16.57, -7.32]. Yüzde 95 GA'nın yorumu P değerinin yorumundan çok daha kolaydır. Örneğin, üçüncü güven aralığı bize CT tedavisinin plaseboya göre ortalama WOMACPA'da gerçekten de anlamlı düzeyde fazla bir düşüş sağlandığını söyler ve gerçek değer (Δ), % 95 ihtimalle, -7.32 ve -16.57 arasında yer almaktadır.

Dahası, % 95 GA ile P değeri arasında bir ilişki vardır. Şöyle ki, Δ 'nın % 95 GA o sayısını içermediği zaman, buna uyan P değeri 0.05'ten azdır veya tersi geçerlidir. Bu nedenle yukarıdaki % 95 GA'larından dolayı, ET plaseboya (Δ 'nın % 95 GA o'ı içermez) göre istatistiksel olarak anlamlı derecede farklı (ve daha iyi) iken, CE ve ET tedavilerinin istatistiksel olarak anlamlı bir farklarının (Δ 'nın % 95 GA o içerir) olmadığına karar verebiliriz.

% 95 GA'nın P değerinden daha çok aydınlatıcı olduğunu anlamamız önemlidir.

Gerçekten her iki metot da ortalamalardaki farkın ve istatistiksel anlamlılığın klinik ilişkisinin değerlendirilmesine izin verir. Ama güven aralığının yorumu çok daha kolaydır ve sonuçların farklı klinik çalışmalarda daha iyi karşılaştırılmasına olanak sağlar. % 95 GA aynı zamanda metaanalizlerde de tercih edilen istatistiksel değerlendirme aracıdır.

Özetle; P değerinin ortaya atılmasından neredeyse bir asır sonra bile çoğu klinisyen tarafından hala yeterince anlaşılmadığını görmek şaşırtıcıdır. Bu nedenle, P değerinin çok özel bir kavram olduğu ve klinisyen tarafından iyi anlaşılmadığı bilinmelidir. Diğer taraftan güven aralığı ise gözlenen klinik sonuçlarda daha iyi bir yorum sağlar ve anlaşılması daha kolaydır. Artan sayıda yayının artık güven aralığı kullanmaya başlaması memnuniyet vericidir.

REFERANSLAR

1. C. O. Bingham III, A. I. Sebba, B. R. Rubin, G. E. Ruoff, J. Kremer, S. Bird, S. S. Smugar, B. J. Fitzgerald, K. O'Brien and A. M. Tershakovec. Efficacy and safety of etoricoxib 30 mg and celecoxib 200 mg in the treatment of osteoarthritis in two identically designed, randomized, placebo-controlled, non-inferiority studies, *Rheumatology* 2007;436:496-507
2. Lesaffre, E., Superiority, equivalence and non-inferiority trials. *Bull NYU Hospital Jt Dis* 2008; 66 (2):150-4
3. Moyé, L.A., *Multiple Analyses In Clinical Trials: Fundamentals for Investigators*, Springer - Verlag, New York, 2003
4. Proschan, M.A., Lan, K.K.G. and Wittes, J.T. *Statistical Monitoring of Clinical Trials: A Unified Approach*, Springer - Verlag, New York, 2006

GİRİŞ

Çalışmanın amacı deneysel (E) tedavinin kontrol (C) tedavisine üstün olduğunu göstermekse, randomize kontrollü çalışma üstünlük çalışması diye adlandırılır. İlgili istatistiksel test ise üstünlük testidir. Anlamlı bir sonuç elde edilen bir üstünlük çalışmasında E'nin C'ye göre istatistiksel açıdan anlamlı düzeyde farklı bir etkisi olduğuna karar verilebilir ve gözlenen cevap E'nin tarafındaysa, E'nin istatistiksel olarak anlamlı derecede C'den daha etkili olduğu sonucuna varabiliriz. Bununla birlikte anlamlı olmayan bir sonuç elde edilen durumda, E'nin C'ye göre daha iyi olduğunu kimse iddia edemez ama yine kimse E'nin C kadar iyi olduğunu da iddia edemez. Aslında E ve C aynı tedavi olmadığı sürece, etki bakımından her zaman küçük bir fark olacaktır ve her bir küçük gerçek fark (Δ) için eşit etki, yokluk hipotezinin (*null-hypothesis*) yüksek olasılıkla reddedilebileceği bir örneklem büyüklüğü kurulabilir (bkz. Lesaffre (2008)). Anlamlı olmayan bir sonuç, maalesef, sıklıkla kanıtın yokluğu şeklinde yanlış olarak yorumlanır.

Günümüzde ilaç firmaları aynı etkinliğe sahip ama diğer özellikler bakımından daha iyi kalitede yeni tedaviler aramaktadır. Bu eğilim, deneysel tedavinin kontrol tedaviye göre daha az etkili olmadığını göstermeyi amaçlayan randomize kontrollü çalışmaların ortaya çıkmasını artırmıştır. Bu tür çalışmalar *daha az etkili olmama* çalışmaları olarak adlandırılmaktadır.

Konuşmam, üstünlük, eşdeğerlik ve *daha az etkili olmama* çalışmalarının temel felsefelerini karşılaştırmayı amaçlamaktadır. Dahası, bir *daha az etkili olmama* çalışmasının kurgulanması ve yürütülmesine ilişkin pratik problemler tartışılacaktır. Kavramlar trombolitik ilaçlarda kurgusal örnekler yoluyla açıklanacak ve bazı örnekler aynı zamanda Bingham ve arkadaşlarının (2007) çalışmasından alınacaktır. Lesaffre (2008) özellikle etorikoksib 300 mg qd (ET) ve selekoksib 200 mg qd (CE)'yi plaseboyla (PL) karşılaştıran iki çalışmanın ilk bölümünü örnek almıştır. Çalışmanın ikinci bölümünde, ET ve CE'nin performanslarını *daha az etkili olmama* tasarımı ile karşılaştırmak üzere iki çalışma yürütülmüştür. Burada, çalışmanın özellikle *daha az etkili olmama* kısmı tartışılacaktır. Çalışma ile ilgili ayrıntılı bilgi için

ÜSTÜNLÜK, EŞDEĞER ETKİNLİK VE DAHA AZ ETKİLİ OLMAMA ÇALIŞMALARI (Superiority, equivalence and non-inferiority)

Emmanuel Lesaffre

Leuven Katolik Üniversitesi,
Belçika
Erasmus Tıp Merkezi, Rotterdam,
Hollanda

Lesaffre (2008) ve orijinal yayına bakınız.

ÜSTÜNLÜK ÇALIŞMASI

Örnekleri trombolitik ilaçlarla akut MI tedavisi gibi özellikle uzmanlık alanım olan bir konudan seçtim. Bu terapötik alanda *daha az etkili olmama* tasarımları on yıldan uzun bir süre önce başlatılmıştır. Önce GUSTO-I çalışmasına (The Gusto Investigators, 1993) bakalım. Bu çalışmada dört kol bulunur ama biz burada sadece tedavi kollarını göz önüne alıyoruz: Streptokinaz + intravenöz heparine (C) karşılık rtPA + intravenöz heparin (E). Primer sonlanım noktası, 30 günlük mortalitedir. Akut MI'dan sonraki 30 gün içinde hasta öldüyse ikili (binary) sonuç 1'e eşittir veya diğer deyişle sıfırdır. Çalışmada 10,370 hastaya kontrol ve 10,348 hastaya da deneysel tedavi uygulanmıştır: 30 gün içinde ölen hastaların gözlenen yüzdesi kontrol tedavisi için % 7.4 ve deneysel tedavi için % 6.3'tür. $H_0: \Delta = 0$ 'ı olarak değerlendiren ki-kare testinin P değeri 0.0028 olup, Δ için iki yönlü % 95 GA değeri % 0.36-%1.73 arasındadır. Sonuç, rtPA'nın streptokinaza göre istatistiksel açıdan anlamlı derecede daha düşük 30 günlük mortalite oranına sahip olduğunu gösterir. Aynı sonucun 1000 hastanın iki katı ile de elde edildiğini varsayarsak, $P=0.37$ ve Δ % 95 GA [-% 1.21, % 3.21] olur. Bu durumda; iki tedavinin eşit etkinliğe sahip olduğu sonucuna mı varıyoruz? Tabii ki daha büyük örneklem büyüklüğü için fark yaratmamız mümkün ama toplam örneklem büyüklüğü 2000 hasta olduğunda gruplar arasında açık bir ayrımın görülməsi mümkün olmamıştır: Lesaffre (2008)'de açıklandığı gibi, farklı bir tedavi etkisi için kanıt yoktur ama farklı bir tedavi etkisinin olmaması için de hiçbir kanıt yoktur.

İki yönlü üstünlük çalışması için geçersiz ve alternatif hipotezlerin sırasıyla $H_0: \Delta = 0$ ve $H_A: \Delta \neq 0$ olduğunu hatırlayınız (Lesaffre, 2008'e de bakınız). İstatistiksel test teorisinde çok iyi bilinen aşağıdaki özelliği kullanarak üstünlük, eşdeğerlik ve *daha az etkili olmama* istatistiksel özelliklerinin karşılaştırılması yararlı olacaktır (örneğin Lesaffre, 2008'e bakınız). Eşleştirilmemiş t-testi ile iki ortalama karşılaştırılırken, 0.05 ($P < 0.05$)'te anlamlı bir sonuç, ortalamaların farkı için Δ ya ait % 95 GA'nın 0'ı içermesine eşdeğerdir. Diğer taraftan, anlamlı olmayan

sonuç ($P > 0.05$), % 95 GA'nın 0'ı içermemesi sonucu ile eşdeğer olur. İki oranın karşılaştırılması için bu sonuç sadece yaklaşıktır ancak pratikte eşdeğerlik neredeyse her zaman geçerlidir. Şekil 1'de E 'nin C 'den $\alpha = 0.05$ 'te üstün olduğunu ima eden, 30 günlük mortalite oranındaki fark için % 95 GA'nın sıfırı kapsamadığı bir randomize kontrollü çalışmanın sonuçları gösterilmektedir. Yayının geri kalanında üstünlük çalışması için örneklem büyüklüğünün hesaplanmasında kullanılan varsayılan Δ , Δ_S olarak gösterilecektir.

EŞDEĞERLİK ÇALIŞMASI

Yukarıdaki kurgusal örnekle devam ediyoruz. Randomize kontrollü çalışmaların E ve C tedavilerinin eşit etkinlikte olduğunu göstermeyi amaçladığını varsayalım. Lesaffre (2008)'den biliyoruz ki, bunu istatistiksel testlerle göstermemiz imkansızdır; dolayısıyla, "eşit iyi" için pratik bir tanıma yönelmemiz gerekir. Gerçek fark Δ 'nın klinik eşdeğerlik aralığının $[-\Delta_E, \Delta_E]$ içine denk düştüğü zaman iki tedavinin (çok da fazla) fark etmediği düşünülebiyecek bir Δ_E değeri üzerinde klinisyenlerin anlaştıklarını varsayalım. Eğer gözlenen fark ve % 95 GA tamamen klinik eşdeğerlik aralığının içine denk düşerse, iki tedavi "eşdeğer" olarak adlandırılabilir. Yokluk hipotezi ya da alternatif hipotez koşullarında eşdeğerliğin kanıtlanması, $H_0: \Delta > \Delta_E$ veya $\Delta < -\Delta_E$ 'nin, alternatif hipotez $H_A: \Delta_E < \Delta < -\Delta_E$ yönünde, uygun bir istatistiksel test ile reddetmek şeklinde daralacaktır. Oranlar karşılaştırılırken ki-kare testinin daha basit bir adaptasyonuna ihtiyaç vardır. Dahası, bir eşdeğerlik çalışması için anlamlı sonuç ($P < 0.05$), klinik olarak tanımlanan eşdeğerlik tanımına göre iki tedavinin eşdeğer olmasıdır. Şekil 1'de, Δ_E 'nin 0.01 (% 1) olarak tanımlandığı kurgusal bir eşdeğerlik çalışmasından elde edilen anlamlı sonuç gösterilmektedir. Klinik eşdeğerlik aralığının bir veya her iki sınır değerlerini kesen % 95 GA'na karşılık gelen, $P > 0.05$ durumunda, iki tedavi eşdeğer olarak adlandırılmaz.

Eşdeğerlik çalışmalarının çoğu, en yaygın farmakokinetik değişken C_{max} , C_{min} ve AUC ile ifade edilen, "aynı" farmakokinetik profile sahip olduğunu göstermek için, jenerik bir ilacın orijinal ticari bir ilaçla karşılaştırılmasını amaçlayan biyoeşdeğerlik çalışmalarıdır. Δ_E 'nin, "bir ilacın diğeri ile değiştirildiği zaman hastada hiçbir fark

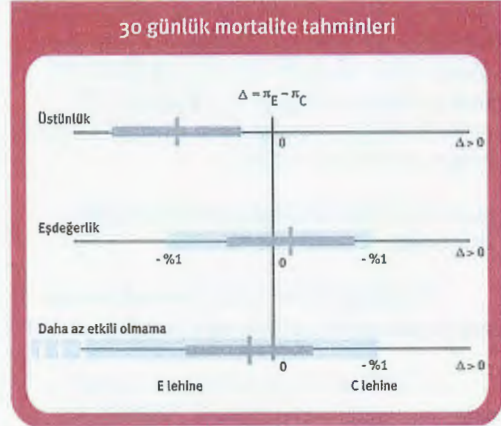
tespit edilemeyeceği” değeri olarak tanımlanabildiğine dikkat ediniz. Dahası, biyoeşdeğerlik çalışmalarının % 95 GA yerine sıklıkla % 90 GA’nı kullandığını söylemek gereklidir. *Daha az etkili olmama* çalışmalarından (sonraki bölüm) bazen (yanlışlıkla) eşdeğerlik çalışması olarak bahsedildiği gözlenir.

DAHA AZ ETKİLİ OLMAMA ÇALIŞMASI

Eşdeğerlik çalışmaları terapötik çalışmalar için uygun değildir. İkinci kurgusal örnekte Δ için % 95 GA’nın klinik eşdeğerlik aralığının sol sınırını kestiğini varsayalım. Bu durum E’nin C’ye üstün olduğunu veya en çok Δ_E ’nin C’den daha kötü olduğunu ifade eder. Bir *daha az etkili olmama* çalışmasında kimileri sadece üst sınır Δ_{N1} tanımlar ve Δ için % 95 GA’ı Δ_{N1} ’nın soluna denk düştüğü zaman E tedavisinin C’den daha az etkili olmadığı şeklinde adlandırılır. *Daha az etkili olmama* çalışmasında sadece klasik iki yönlü % 95 GA sınırına baktığımızı dikkat çekerim; bu gerçekte bir kenarı sınırsız olan (burada sol tarafta) bir yönlü % 97.5 GA’na karşılık gelir.

Bu bir yönlü güven aralığı, bir yönlü anlamlılık düzeyi 0.025’e uyar. Bu nedenle, üstünlük ve eşdeğerlik testlerinde, kimileri, anlamlılık testini (gerçekte iki tane bir yönlü testlerin bileşimi olan) iki yönlü 0.05 seviyesinde uygularken, standart bir *daha az etkili olmama* testini bir yönlü 0.025 seviyesinde uygular. Dahası, üstünlük çalışmalarında halen sonuçların bildirilmesi için P değeri popülerken, *daha az etkili olmama* çalışmasında (iki yönlü) % 95 GA’nın kullanımı daha sıktır.

Daha az etkili olmama’nın kanıtlanması, $H_0: \Delta > \Delta_{N1}$ ’nin, alternatif hipotez $H_A: \Delta < \Delta_{N1}$ yönünde (bir eşdeğerlik testi için olarak) adapte edilmiş bir klasik istatistiksel test ile reddedilmesi gereğini doğurur. Bu durumda anlamlı bir sonuç ($P < 0.025$), deneysel tedavinin kontrol tedavisine göre daha az etkili olmadığı anlamına gelir ve daha az etkili olmama sonucu seçilen bir Δ_{N1} değerine dayanır. Şekil 1’de, Δ_{N1} ’nin 0.01 (%1) olarak tanımlı olduğu kurgulanmış bir *daha az etkili olmama* çalışmasının anlamlı sonucu gösterilmektedir. $P > 0.025$ durumunda, örneğin bir kenarlı % 97.5 GA’nın (iki kenarlı % 95 GA) sağ sınırı Δ_{N1} ’yı aştığı zaman, deneysel tedavinin daha az etkili olmadığını göstermekte başarısız



Şekil 1: Üstünlük, eşdeğerlik ve daha az etkili olmama kavramları

oluruz. Bununla birlikte bu durum, tedavi C'nin E'ye üstün olduğunu gösterdiğimiz anlamına da gelmez. *Daha az etkili olmama* aralığının sol tarafta sınırsız olduğuna dikkat ediniz. Bu nedenle de bu bölgeden *daha az etkili olmama* bölgesi olarak bahsedilir.

DAHA AZ ETKİLİ OLMAMA ÇALIŞMASINA İKİ ÖRNEK

1. Etorikoksib 300 mg ve selekoksib 200 mg'ın osteoartrit tedavisinde karşılaştırılması

Herbiri Bingham ve arkadaşları (2007) tarafından yapılan çalışmada (a) WOMAC Ağrı Alt-ölçeği (WOMACPA), (b) WOMAC Fiziksel Fonksiyon Alt-ölçeği (WOMACPH) ve (c) Hastalık Durumunun Hasta Global Değerlendirmesi (PGADS)'nin başlangıçtan 12. haftaya kadar, zaman ağırlıklı olarak değişikliği kullanarak diz ve kalçalardaki osteoartrit tedavisi için her biri CE'ye karşılık ET'nin *daha az etkili olmama* değerlendirmesi içeren iki adet, üç kollu randomize klinik çalışma söz konusudur. Ayrıntılı bilgi orijinal yayında veya Lesaffre (2008)'de bulunabilir.

CE ve ET arasındaki fark için iki yönlü % 95 GA'larının üst sınırı, üç primer son nokta WOMACPA, WOMACPH ve PGADS için en çok 10mm olduğu zaman "deneysel tedavi CE, ET'den daha az etkili değildir" olarak verilmiştir. Bundan dolayı, *daha az etkili olmamanın* gösterilebilmesi için her üç koşulun sağlanması gerekmektedir. Bingham ve arkadaşlarının (2007) çalışmasındaki Tablo 2a ve 2b, iki çalışma için bu koşulların (% 95 GA'ları tamamen üst sınır altındadır) sağlandığını ve yazarın sonucunun bu nedenle "etorikoksib 300 mg'ın osteoartritte celekoksib 200 mg'a kıyaslanabilir" şeklinde olduğunu göstermektedir. İlk bakışta, yazarlar "*daha az etkili olmama*" için zor bir kriter kullanmıştır, ancak $\Delta_{NI} = 10\text{mm}$ 'nin nasıl seçildiği belirgin değildir.

Son olarak Tablo 2a ve 2b'de sadece % 95 GA'ları değil, P değerleri de bildirilmektedir. Bununla birlikte, CE ile ET'yi karşılaştıran tüm P değerleri anlamlı değildir. Şunu gözden kaçırmamak önemlidir; bu P değerleri, CE ile ET'yi üç sonlanım noktasında karşılaştıran üstünlük testlerine tekabül eder. Bunun sonucu olarak da, yukarıda açıklanan yokluk ve alternatif

daha az etkili olmama hipotezlerine ait değildir.

2. ASSENT II çalışması

ASSENT II (ASSENT-II Araştırmacıları, 1999), akut MI hastalarının tedavisi için trombolitik tedavinin yeriyile ilgili ilk *daha az etkili olmama* çalışmalarından biridir. ASSENT II'de, Boehringer Ingelheim tarafından üretilen tek bolus dozlu tenekteplaz (E) ile alteplaz (C) hızlandırılmış infüzyonundan sonraki 30 günlük mortalite oranlarını karşılaştırılır.

Çalışmada önce, *daha az etkili olmama* alanının üst sınırı, Δ_{NI} , 0.01 (% 1) olarak alınmıştır. Bu, E'nin C'ye göre (alan içindeki % 95 GA ile) % 1 daha yüksek 30 günlük mortalite oranı gösterdiğini ifade eder. Bununla birlikte, kontrol tedavisi için gerçek 30 günlük mortalite oranı (π_C)'nda belirsizlik vardır. Geçmiş çalışmalara dayanarak, π_C için gerçek değer muhtemelen 0.05 (% 5) ve 0.10 (% 10) arasına denk düştüğü sonucuna varılabilir. Göreceli koşullarda, gerçek oran % 5 olduğu zaman E için % 1 daha kötü performans oranı, gerçek kontrol oranı % 10 olduğu zaman % 10'a kıyasla % 20'dir. Bu nedenle, π_C için küçük değerler olan $\Delta_{NI} = 0.01$, kabul edilemez olarak düşünülmüştür ve ikinci bir adımda *daha az etkili olmama* alanı $\pi_C > 0.072$: $\Delta = \pi_E - \pi_C < 0.01$ ve $\pi_C \leq 0.072$: $\pi_E/\pi_C < 1.14$ olarak değiştirilmiştir. Burada iki kriterin $\pi_C = 0.072$ bakımından eşit olduğuna dikkat ediniz. Kesme noktası 0.072'nin seçimi için, ASSENT-II Araştırmacıları referans gösteriyoruz (1999); 0.01 ve 1.14 değerleri tarihsel çalışmalardan (GUSTO III) ve bir metaanalizden alınmıştır. O günlerde, tek yönlü % 95 GA (iki kenarlı % 90 GA)'nın tanımı için kullanılabildiğinden, *daha az etkili olmama* kriterinin daha az katı olduğu gözlenir.

Primer sonlanım noktası sonuçları, E (tenekteplaz) için % 6.16 ve C (alteplaz) için de % 6.18'dir. Gözlenen 30 günlük mortalite oranı % 7.2'nin altındadır. Protokol, rölatif kriterin bu durumda kullanılması gerektiğini belirtmiştir. Gözlenen relatif risk her ikisi de 1.14'ten küçük olmak üzere % 90 GA = [0.904-1.101] ile 0.997 idi. Sonuç bu nedenle, tenekteplazın alteplazdan daha az etkili olmadığı yönündedir.

DAHA AZ ETKİLİ OLMAMA ÇALIŞMALARININ ELEŞTİREL TARTIŞMASI

1. Daha az etkili olmama tasarımı seçme nedenleri

Deneyssel tedavi E'nin kontrol tedavisi C'den daha az etkili olmamasının gösterilmesi aşağıdaki nedenlerden dolayı ilginç olabilir:

- *Etik anlamda plasebo kontrollü çalışma yapmak artık mümkün değildir.* Streptokinazın akut MI hastalarında 30 günlük mortaliteyi önemli derecede düşürdüğü gösterildiğinden, alteplazın piyasaya verildiği zaman (GUSTO Araştırmacıları, 1993) performansının plaseboya göre karşılaştırılması mümkün değildi. O sıralarda, yine de deneyssel tedavinin daha iyi olması gerektiğine (bu vakada öyle olduğu kanıtlanmıştı da) inanılıyordu. ASSENT II çalışmasının planlandığı sıralarda, tenekteplaz için bu varsayılmazdı ve dolayısıyla, *daha az etkili olmama* tasarımı çözüm olarak görüldü.
- *Primer etkinlik sonlanım noktasında E'nin C'den daha iyi olması beklenmez ama sekonder sonlanım noktalarında daha iyi veya daha güvenlidir.* Örneğin, selekoksib analjezik ve antiinflamatuvar etkili bir non-steroidal antiinflamatuvar ilaçtır ama muhtemelen ciddi gastrointestinal etkilere neden olmaktadır. Tipik bir osteoartrit hastası, potansiyel ilişkili faktörler ve NSAID'lerin uzun süreyle yüksek dozdaki kullanımı nedeniyle NSAID gastropatisi açısından yüksek risk taşır. Bununla birlikte klinik çalışmalarda, etorikoksib gibi COX-2 inhibitörleri, osteoartrit ağrı tedavisinde NSAID'lerle benzer etkinliğe ama daha az gastrointestinal yan etkilere sahiptir. Bir *daha az etkili olmama* tasarımı bu nedenle uygundur (Bingham ve arkadaşları 2007).
- *Primer etkinlik sonlanım noktasında, E'nin C'den daha iyi olması beklenmez ama üretimi daha ucuz veya uygulanması daha kolaydır.* ASSENT II çalışması için 90 dakika (hızlandırılmış) infüzyon gereken alteplaz uygulamasından kesinlikle daha kolay olarak, tenekteplazın uygulanması için tek bolus doz enjeksiyonu gerekir.
- *Klinik çalışmada primer etkinlik sonlanım noktasında E'nin C'den daha iyi olması beklenmez, ama klinik çalışma dışı kullanımda hasta uyumu daha iyi olacaktır ve dolayısıyla çalışma dışında etkinlik daha iyi olacaktır.*

Kontrol tedavisinin oral bir tedaviye katı uyum gerektirdiğini varsayalım; örneğin hastaların ilacı düzenli olarak ve dar zaman aralıklarında almasını gerektiren durumlar. Randomize kontrollü bir çalışmada hastalar yakından takip edilir ve önerilen tedaviye uymaları için teşvik edilir; böylece, ilacı düzenli alma oranı göreceli olarak oldukça yüksek olacaktır. Randomize kontrollü çalışma dışında, hastalar daha az disiplinli olacaktır ve bundan dolayı, tedavi etkisinin randomize kontrollü çalışma dışında çok daha düşük olması muhtemeldir. Sonuç olarak, kontrol tedavisi ile neredeyse aynı etkinlikte olan ancak daha kolay uygulanan bir deneyssel tedavi, çalışma dışında muhtemelen kontrol tedavisinden daha yüksek etkinliğe sahip olacaktır.

2. Daha az etkili olmama sınırlarının belirlenmesi

Daha az etkili olmama sınırı Δ_{NI} 'nin seçilmesi için iki yol vardır. İlk olarak, E ve C arasında doğrudan bir karşılaştırma düşünülebilir. Bu durumda, Δ_{NI} sadece klinik temele dayandırılır. Klinisyenlerin etkilerin farklılığını göstereceğini düşündükleri noktayı (kesme noktasını) seçmeleri gerekir. Bununla birlikte, böyle bir kesme noktasının seçimi özellikle hayatı tehdit edici hastalıklarda sıklıkla çok zordur. Örneğin, akut MI hastaları için 30 günlük mortalitede % 1 seçimini nasıl destekleyebiliriz? Deneyssel tedavinin kontrol tedaviye göre daha az yaşamı korumasının bir sorun olmadığını hiçbir klinisyen iddia edemez. Belirli bir terapötik alanda ilk kez bir *daha az etkili olmama* tasarımı kuruyorsanız, bu da bir problemdir. Böyle bir durumda, örnek alınacak bir değer yoktur ve FDA ve/veya EMEA ile kabul edilebilir bir sınır değeri üzerine pazarlık etmeniz gerekebilir. İkinci neden ise, plasebo tedavisi ile dolaylı yoldan karşılaştırmaya dayanır. Böyle bir Δ_{NI} durumunda C'den daha az etkili olmasına izin verilse de, E'nin plaseboya üstünlüğünü garanti edecek istatistik tartışmalara dayanılarak belirlenir. İkinci yaklaşım klinik olarak kabul edilebilir ve E'ni plaseboya üstünlüğünü garanti edecek bir Δ_{NI} ile sonuçlanacak şekilde, klinik ve istatistiksel görüşlerin birleştirilmesiyle düzenlenebilir.

İkinci yaklaşım, kontrol tedavisi ve plasebo arasındaki muhtemel etkinlik farkının tahmin edilmesini gerektirir. Bu, örneğin bir metaanalizden

elde edilebilir. Bu farkı % 95 GA [% 1.7-2.3] ile % 2 olarak bulduğumuzu varsayalım. Kontrol tedavisinin plaseboya karşılık farklı etkinliği üzerine belirsizliği göz önüne alarak, deneysel tedavinin plaseboda daha iyi olduğunu (% 95 güven ile) garanti etmek için, deneysel tedavi kontrol tedavisinden en çok % 1.7 daha kötü olabilir ve dolayısıyla $\Delta_{NI} < \% 1.7$ 'dir. Yüzde 1.7'nin klinik olarak kabul edilebilirliği veya ileri seviyede düşüşün gerekli olup olmadığı klinisyenlerce kontrol edilmelidir. Kontrol ve plasebo tedavisi arasında tahmin edilen minimal farkın yarısını almak popüler bir kuraldır: örneğin burada, $\% 1.7/2 = \% 0.85$ olacaktır.

Bu ikinci yaklaşım, kontrol tedavisinin çok iyi tanımlanmış, tahmin edilebilir ve ölçülebilir bir etkisi bulunmasını gerektirir. Bu durum, çok sayıda plasebo kontrollü randomize çalışmaların daha evvel yapılmış olmasını gerektirir. Eğer durum böyle değilse, deneysel tedavinin plaseboda daha iyi olduğu yönünde "kanıtlanamaması" gibi bir risk her zaman vardır. Aynı zamanda, kontrol tedavisinin göreceli olarak plaseboda daha iyi sabit performans göstermesi de gerekmektedir (sabitlik varsayımı). Aksi durumda, Δ_{NI} için doğru değerin seçimi zor olacaktır. Kontrole karşı plasebo tedavisinin etkisinin doğru tahmin edilmemesi problemi maalesef ki zaman içinde plasebo tanımının dramatik olarak değişebilmesi nedeniyle her zaman problem olarak kalacaktır.

Yukarıdaki *daha az etkili olmama* çalışmalarına göre, ikinci yaklaşım ASSENT II *daha az etkili olmama* alanının tayini için kullanılırken, Δ_{NI} 'nin osteoartrit çalışması için tamamen klinik temele göre seçildiği görülmektedir.

3. Intention-To-Treat (tedaviye niyete göre) mi yoksa protokole göre analiz mi?

Bir üstünlük çalışması için önerilen hasta popülasyonu, *Intention-To-Treat (ITT)* popülasyonudur. Bu, randomize edilen tüm hastaları ve bu nedenle de tedavi edilmesi beklenen tüm hastaları içerir. Protokolü doğru kullanmayanlar, bir veya daha çok kontrol vizitini kaçıran hastalar, çalışma dışı kalan hastalar, yanlış gruba randomize edilen hastalar vs. planlanan tedaviye göre analiz edilirler. ITT

yaklaşımı ideal olmamakla birlikte, ITT prensibi çalışmanın sonucu üzerine konservatif bir etki oluşturabileceğinden, üstünlük çalışmaları için en uygun yaklaşım olduğu düşünülmür. Şöyle ki, çalışma zayıf yürütüldüğünde, deneysel tedavinin kontrol tedavisinden daha etkin olduğunun kanıtlanması pek olası olmayacaktır.

Daha az etkili olmama çalışması için ise ITT analizinin konservatif bir etkisi yoktur. Çalışmadan ayrılmalar ve çalışmanın zayıf yürütülmesi, iki kolun sonuçlarının birbirine yönelmesine neden olabilir. Diğer bir olasılık ise, sadece protokol dışı davrananlardan oluşan *per-protokol (PP)* popülasyona başvurmaktır. Bununla birlikte, PP analizinin bir *daha az etkili olmama* tasarımı için yeterli konservatif etkisinin olup olmadığı açık değildir. Bu probleme iyi bir çözüm mevcut olmadığından, *daha az etkili olmama* çalışması için önerilen pragmatik yaklaşım her iki analizin de (ITT ve PP) uygulanması şeklindedir. İki analizin birbirini doğrulayacağı düşünülmür.

4. Örneklem büyüklüğü hesaplamaları

Bir *daha az etkili olmama* çalışması için gerekli olan örneklem büyüklüğü, n , başka faktörlerin yanı sıra klinik olarak önemli fark olan Δ_S 'ye bağlıdır. Bir *daha az etkili olmama* çalışması için gerekli olan örneklem büyüklüğü, *daha az etkili olmamanın* üst sınırı olan, Δ_{NI} 'ye bağlıdır. $\Delta_S = \Delta_{NI}$ olduğunda, $\Delta = 0$ varsayımı altında gerekli örneklem büyüklüğü buna tekabül eden üstünlük çalışma testi $\Delta = 0$ için örneklem büyüklüğüne ve eğer $\Delta = \Delta_S$ ise eşittir. Diğer taraftan, *daha az etkili olmama* çalışması için örneklem büyüklüğünün üstünlük çalışmasıninkinden sıklıkla daha büyük olmasına neden olacak şekilde, Δ_S tipik olarak Δ_{NI} 'den (çok daha fazla) büyüktür.

5. Daha az etkili olmama ve üstünlüğün bir çalışmada birleştirilmesi

Daha az etkili olmama planlandığı ve çalışma sonunda gerçekleştiği zaman, daha ileri gidip çok sayıda anlamlılık testi açısından için istatistiksel bir hata olmaksızın aynı verilerle üstünlüğün de kanıtlanıp kanıtlanmayacağı veya üstünlüğün kanıtlanması başarısız olursa, aynı verilerle istatistiksel bir hata olmaksızın *daha az*

etkili olmamanın gösterilip gösterilemeyeceği sorulabilir.

Söyle ki (bkz. örneğin Moyé, 2003): aynı popülasyona uygulanmış (ITT veya PP) *daha az etkili olmama* ve üstünlük, “Kapalı Test Prensibi” nedeniyle istatistiksel bir hata olmaksızın 0.05’te sınılanabilir (bkz. Lesaffre (2008)). *Daha az etkili olmama* ilk olarak ITT popülasyonunda sonra da üstünlük ITT popülasyonunda test edildiği zaman, yine bunun olumsuz bir sonucu olmaz. Ancak bunun tersini yapmak çok sayıda test için düzeltme yapmayı gerektirir.

Özetle, üstünlüğü göstermeyi amaçlayan randomize kontrollü çalışma, klinik çalışma araştırmalarında altın standarttır. Üç tip çalışmadan üstünlük çalışmasını yorumlamak en kolaydır. Daha güçlü ilaçların piyasaya verilmesindeki zorluklar nedeniyle, ilaç endüstrisi mevcut en etkin ilaçları geliştiremeyen ama tedavinin diğer özellikleri bakımından daha iyi olacağını düşündükleri ilaçları aramaya zorlanmıştır. Bu durum hastanın yaşam kalitesini iyileştirdiği sürece, bu tür gelişmeler sevindiricidir. Bununla birlikte klinik araştırmacı için, *daha az etkili olmama* çalışmalarının uygulaması ve bazen de yorumlanması pek kolay değildir. İlkokulda matematik dersinde A, B ve C sayıları ile ilgili aşağıdaki kural öğretilmiştir: (a) eğer $A = B$ ve $B = C$ ise $A = C$ ve (b) eğer $A < B$ ve $B < C$ ise $A < C$. Bununla birlikte, ya “ $\approx$ ” işareti “eşdeğer” ve “ni” *daha az etkili olmama* (non-inferior) demek ise, o zaman aşağıdakileri söyleyebilir miyiz:

- Eğer $A \approx B$ ve $B \approx C$ ise, $A \approx C$ midir?
- Eğer $A \neq B$ ve $B \neq C$ ise, $A \neq C$ midir?
- Eğer $A \neq B$ ve $B \neq P$ ise, $A \neq P$ midir?

Bu soruların cevabı belirgin değildir ve sırasıyla eşdeğerlik aralığı ve *daha az etkili olmama* alanının seçimine bağlı olacaktır. Eşdeğerlik ve *daha az etkili olmama* için farklı tanımlar, klinik araştırmacılar ve klinik karar vericilerin hayatını zorlaştırarak, bir *daha az etkili olmama* çalışmasının gerçekte ne mesaj taşıdığına karar vermelerini zorlaştıracaktır. Sonuç olarak, Le Henanff ve arkadaşları (2006) 1

Ocak 2003 ve 31 Aralık 2004 arasında yayımlanan *daha az etkili olmama* ve eşdeğerlik çalışmalarını gözden geçirmişlerdir ve bu tür çalışmaların önemli eksiklikleri olduğu sonucuna varmışlardır.

REFERANSLAR

1. ASSENT-2 Investigators. Single-bonus tenecteplase

- compared with front-loaded alteplase in acute myocardial infarction: the ASSENT-2 double-blind randomised trial. *The Lancet*, 1999, 354, 716-722.
2. C. O. Bingham III, A. I. Sebba, B. R. Rubin, G. E. Ruoff, J. Kremer, S. Bird, S. S. Smugar, B. J. Fitzgerald, K. O'Brien and A. M. Tershakovec. Efficacy and safety of etoricoxib 30 mg and celecoxib 200 mg in the treatment of osteoarthritis in two identically designed, randomized, placebo-controlled, non-inferiority studies, *Rheumatology* 2007;46:496-507
 3. Le Henanff, A., Giraudeau, B., Baron, G. And Ravaud, P. Quality of reporting of noninferiority and equivalence randomized trials. *JAMA*, 2006;295, 10, 1147-1151.
 4. Lesaffre, E., Use and misuse of the P-value, 2008
 5. Lesaffre, E. Bluhmki, E. Wang-Clow, F. Berlioli, S. Danays, T. Fox, N.L. and Van de Werf, F.. The general concepts of an equivalence trial, applied to ASSENT-2, a large-scale mortality study comparing two fibrinolytic agents in acute myocardial infarction. *European Heart Journal*, 2000, 21(Review Article), 1-5.
 6. Moyé, L.A., *Multiple Analyses In Clinical Trials: Fundamentals for Investigators*, Springer-Verlag New York, 2003
 7. The GUSTO Investigators. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. *New England Journal of Medicine*, 1993, 329, 673-682.

RANDOMİZE KONTROLLÜ ÇALIŞMALARIN DOĞRU VE YANLIŞ KULLANIMI

Hasan Yazıcı

İlk olarak randomize kontrollü çalışmanın (RKÇ) bilimsel bir araç olarak niteliklerinden bahsedeceğim, sonra da günümüzde nasıl kullanılıyor ve ne yapılmalı onlara değineceğim.

Bilim tarihi açısından bakıldığında RKÇ'nin ortaya çıkması II. Dünya Savaşı'ndan hemen sonradır. Anımsanacağı üzere o devir tümünden gelen (*deductive*) araştırma yönteminin dünya bilimine egemen olduğu, Einstein, Feynman, Medawar ve benzerlerinin yıllarıdır ve bu akımın felsefecisi de ünlü K. Popper'dır. Yirminci yüzyılın, bir yerde bugüne dek yapılan tüm bilimsel buluşların çok önemli bir bölümünün, yöntemi, tündengelen, varsayım çürütmeye dayanan yöntemdir. Varsayım ancak çürütülemezse kanıtlanır. İşte RKÇ'da, kanımca bu çok mükemmel bilim aracında da aynen böyle bir varsayım çürütme söz konusudur. RKÇ'da yaptığımız yeni bir tedavi yönteminin, ilacın, geleneksel yöntemle veya plaseboya üstün olup olmadığını sinamaktır. Varsayım yeni ilacın eski ilaca veya plaseboya karşı *üstün olmadığıdır*. Yani RKÇ'da da tüm amaç bir varsayımı çürütmedir ve bunu bilmek fevkalade önemlidir. İlk başta, A ilacının B tedavisine üstün olduğuna dair kanımız vardır. Ama varsayımımız onun üstün olmadığı şeklinde olduğundan çalışmamıza kontrol grupları ekliyoruz, kör gözlem yapıyor, verilerimizi ayrıntılı istatistik irdelemekten ve acımasız eşdeğerlendirmeden geçiyoruz. Eğer bütün bunlar sonucu varsayımımız gerçekten çürütülemezse o zaman yeni ilacın kontrol grubuna göre daha etkili olduğu sonucuna varıyoruz.

Aynı bilimsel yaklaşım yan etkiler için de geçerli. Bilimsel yöntem açısından bakıldığında orada da bir varsayım çürütme işlevi söz konusu. Bu kez deneyimizin kurgusu, çürütmek istediğimiz varsayım, söz konusu yeni ilacın, geleneksel tedavi veya plaseboya karşı daha fazla yan etkisi olduğu. Bunu dürüst bir çaba sonucu gösteremezsek o zaman yeni ilacın yan etkisi artmamıştır diyebiliyoruz.

Size bir örnek vereyim. Geçen yıl bir toplantıya katıldım. O toplantıda bir koksibin kardiyovasküler etkilerinin pek fazla olmadığını göstermek amacıyla 20 bin kişilik bir çalışma yapılması planlanıyordu. Anlatmaya çalıştım.

*İstanbul Üniversitesi,
Cerrahpaşa Tıp Fakültesi, İstanbul*

Bunu bilimsel olarak yapabilmek için çalışmaya hipertansiyonlu, miyokard enfarktüsü geçirmiş veya çeşitli nedenlerle koroner arter hastalığı riski artmış denekleri almak gerek. Ancak onlar arasında riskin artmış olduğu bilimsel olarak sınanabilir. Doğal olarak da böyle bir girişim etik olarak kabul edilebilir değildir.

Anımsayalım. Birkaç yıl evvelinin tüm dünyada büyük yankı uyandıran rofecoxib'in üretici firma tarafından piyasadan çekilmesinden sonra aynı firmanın araştırma laboratuvarlarının başkanı bir açıklama yaptı ve özetle yan etkilerin ortaya çıkmasında en güvenli kaynağın RKÇ'lar olduğunu söyledi. Hiç katılmıyorum. Söz konusu çalışmayı hatırlayalım. Amaç rofecoxib'in kolon adenomu oluşumunu (dolayısıyla kanseri) engellediğini sınamaktı. Bir grup denek rofecoxib diğer bir grup plasebo alırken, bağımsız çalışma deneticileri çalışma başladıktan bir süre sonra rofecoxib alan grupta daha fazla kardiyovasküler olay gördüler ve çalışma durduruldu. Yani olay esasında tesadüfen ortaya çıkmıştı. Eğer bu çalışmayla amaç söylendiği gibi kardiyovasküler olay araştırmak olsaydı planın çalışmanın başından böyle yapılmış olması gerekirdi. Az önce verdiğim koksib örneğinde hastalara ilacın miyokard enfarktüsü yapma olasılığı söylenerek onlardan aydınlatılmış olur almak olabilir mi?

Burada FDA'yi de biraz eleştirmek istiyorum. FDA'nın önlerine gelen ilaç çalışmalarında son yıllarda "yan etki"den çok "etki"ye önem verdiğini gözlemliyorum. Örneğin bir ilaç çalışmasında yan etkinin ortaya çıkma zamanı görülen yan etkinin ilaca bağlı olup olmadığını saptamada, yani nedensellikte, çok önemli olabilir. Ancak ilgili FDA rehberinde (1) konuya yeterli önemin verildiği kanısında değilim. Bunun yanında aynı rehberde sadece bazı ilaçlarda Faz 4 (yani ilaç piyasaya çıktıktan sonra) yan etki izlemesinin şart olabileceği yönünde bir görüş var ki tabii buna da katılmak olası değil. Doğrusu, kuşkusuz, tüm yeni ilaçlarda piyasaya çıkma sonrası yan etki izlemesi yapmak.

RKÇ'lar, doğaları gereği gerek etki gerekse de yan etki araştırmada, bir takım kısıtlamalar taşırlar.

Şöyle sıralayabiliriz:

1. RKÇ'da genellikle tek bir ilaç tek bir hastalığı olanda sınırlıdır. Ayrıca denekler deney ilacı dışında sınırlı sayıda ilaç kullanırlar. Halbuki gerçek hayatta durum böyle değildir.
2. Süre kısıtlılığı
3. Denek sayısının küçüklüğü
4. Denekler ile gerçek hayatta ilacı kullananlar arasında, 1. maddede değinilenler dışında da önemli farklar vardır. Öyle ki:
 - Evsizler
 - Telefonu olmayanlar
 - İlaç bağımlıları
 - Alkolikler
 - Sözüne güvenilmeyenler
 - Genellikle yan etki fazla ortaya çıkaranlar da RKÇ'lara pek katılmazlar.

Uzatma çalışmaları (*extension studies*): RKÇ'larla ilgili diğer bir sorun son yıllarda ortaya çıktı. İlaç firmaları "uzatma çalışmaları" diye yeni bir grup çalışma türü ürettiler. Örneğin bir çalışmada A ilacı ilk başta deneklere 1 yıl verilecek gibi planlanıyor. Ancak 1 yılın sonunda, ilacın olumlu etkisi de gösterilmişken, hadi çalışmayı 1 yıl, 2 yıl daha sürdürmek gibi bir girişim başlıyor. Bir anti-TNF ajan olan etanerceptle böyle bir çalışmaya çok güzel bir örnek var. Böyle uzatma çalışmalarının ise aşağıda sıraladığım gibi bilimsel/etik olarak da birçok olumsuz yanı var (2):

1. Equipose kavramının ihlali. "Equipose" kavramı herhangi bir bilimsel çalışmada çalışmaya başlarken çalışma sonucunun hangi yönde olacağını pek bilinmemesi demek. Öyle ki ideal dünyada, ortada dürüst bir varsayım sınaması varsa çalışma sonunda olumlu veya olumsuz sonucun ortaya çıkması olasılığı % 50 dolayındadır. Tabii RKÇ'larda, en azından Faz I ve Faz II çalışma sonuçlarından ötürü bu pek böyle olmuyor. Diğer bir deyişle Faz I ve II aşamalarında olumlu sonuçlar gözlenmemiş ilaçlar RKÇ aşamasına gelemiyor. Bunu kabul etmek gerek. Ancak uzatma çalışmalarında bu yanlılık doğal olarak giderek çok belirgin bir hale geliyor.
2. Aydınlatılmış olur sorunu: Uzatma çalışmalarında bir yöntemin diğerine üstünlüğü zaten kanıtlanmış durumda.

Aydınlatılmış olur alırken hastadan “daha az iyi” tedavi yöntemine devam etmesini istemek kanımca pek etik değil. Buradan giderek kanımca bundan böyle tüm ilaç çalışmalarında aydınlatılmış olur’un tam ne dediğini yazarların -en azından elektronik ortamda-açıkça belirtmeleri gerek.

3. Kuvvet hesapları ve “intention to treat” analizleri gibi diğer sorunlar.

Ancak tüm bunlardan belki de daha önemlisi böyle uzatma çalışmalarında yeni bilgi elde edilmemesi. İşte bunlardan yola çıkarak ilaç endüstrisi şu uzatma çalışması işini artık bırakmalı diyorum.

Konuşmamın girişinde RKÇ’nın yan etki sinanması bakımından uygun olmayan bir araç olduğunu belirtmiştim. Bir RKÇ’da gözlenen yan etkiler doğal olarak açıkça bildirilecektir. Bunun yanında bildirilen yan etkinin, ilaç kullanımından ne kadar zaman sonra ortaya çıktığı ise nedensellik bakımından büyük önem taşır. Örneğin anafilaktik ilaç reaksiyonları ilaç kullanımından hemen sonra ortaya çıkarlar. Halbuki sitotoksik ilaçlarla ortaya çıkabilen neoplaziler genellikle uzun süre ilaç kullanımına gerek gösterir.

Gözlemlerimiz RKÇ’larda bildirilen yan etkilerin ortaya çıkış zamanlarının yazarlar tarafından çoğu kez belirtilmediğiydi. Bu önemli konuyu nesnel olarak bir açıklığa kavuşturmak için en iyi bilinen romatoloji ve genel tıp dergilerinde 5 yıl içinde çıkmış 70 RKÇ’da (romatolojik hastalıklarda Cox 2 inhibitörleri ve anti-TNF ilaçları kullanımıyla ilgili) gözlenen yan etkilerin çıkış zamanlarıyla ilgili bir çalışma yaptık (3). Özetle gördük ki, bu en ünlü dergilerde çıkan yayınlarda bile % 70 durumunda yan etkinin ortaya çıkma zamanı belirtilmiyor ve bu, maalesef, ciddi yan etkiler için de doğru.

Girişte vurguladığım gibi RKÇ bilim yöntemi bakımından yan etki göstermeye uygun bir araç değil. Hal böyle iken, son zamanlarda yaptığımız ve henüz yayımlanmamış bir çalışmaya göre anti-TNF ajanlar yapılan ve yine en iyi dergilerde yayımlanan makalelerde söz konusu çalışmaların % 70’i Etki ve Güvenlilik (*Efficacy and Safety*) çalışması diye nitelendiriliyor. Halbuki çalışmalar nesnel olarak incelendiğinde

ancak çok az sayıda çalışmanın bilimsel olarak yan etki çıkmasını irdelleyebileceği ortaya çıkıyor. Bütün bunlardan öte söz konusu çalışmalarda elde edilen çalışmanın denek sayısının azlığı bakımından özellikle ender görülen yan etkileri bilimsel olarak gösteremeyeceği yazarlar tarafından açıklanmıyor. Yine bu çalışmalarda kuvvet (gerekli denek sayısı) hesapları gerçekçi yapılmamış. Örneğin ankilozan spondilitte infliximab kullanımıyla ilgili Almanya’da bir çalışma yapılıyor ve bu çalışma ilacın olumlu etkisini kesin olarak gösteriyor. Ancak bilimsel olarak açıklanması güç nedenlerle birkaç yıl sonra Hollanda’da tekrarlanıyor ve bu çalışmanın kuvvet hesaplarında birinci çalışmada çıkan sonuçlardan pek yararlanılmıyor.

Özetle RKÇ yeni bir tedavi yönteminin (veya ilacın) geleneksel yöntemle üstünlüğünü a. tümdengelen; b. varsayım çürütmeye dayalı yöntemle kanıtlamaya çalışır. RKÇ bu bilimsel özellikleriyle çok daha ön planda bir “etki sinama” aracı olup “yan etki sinama” aracı olmaktan uzaktır. Son yıllarda bu önemli ve güncel tıbbın gelişmesine çok yardımcı olmuş aracı esas amacından saptırma eğilimleri vardır. Böyle eğilimler, en son değerlendirmede, gerek hastalara, gerek hekimlik mesleğine ve gerekse de ilaç endüstrisine zarar verecektir (4).

REFERANSLAR

1. Reviewer Guidance. Conducting a clinical safety review of a new product application and preparing a report on the review. 2005; <http://www.fda.gov/>
2. Taylor WJ, Weatherhall M. What are open – label extension studies for? J Rheumatol 2006;33:642-3.
3. Yazici Y, Yazici H. A survey of inclusion of the time element when reporting adverse effects in randomized controlled trials of cyclooxygenase-2 and tumor necrosis factor alpha inhibitors. Ann Rheum Dis 2007;66:124-7.
4. Yazici H. Use and abuse of the controlled clinical trial. Bull NYU Hosp Jt Dis. 2007;65:132-4.

Giriş

Romatoid artritli (RA) hastalar için longitudinal veri tabanları ve kayıtlarını üç açıdan ele alacağız: 1) hastaları değerlendirmek ve izlemek için kullanılan yollar; 2) RA’te klinik çalışmaların kısıtlamaları ve longitudinal veri tabanlarına neden gerek var; 3) klinik romatologlar için kendi hastalarına dair kantitatif veri toplama stratejileri.

A. ROMATOİD ARTRİTİ (RA) DEĞERLENDİRME YOLLARI

RA’te kantitatif değerlendirme pek çok klinik durumun ölçümünden farklıdır, zira klinik deneylerde, klinik araştırmalarda ve klinik bakımda münferit vakalara uygulanabilecek kan basıncı, serum glikozu, nabız vs. gibi tek bir “altın standart” yoktur. RA’li hastaları değerlendirmek için kullanılan ölçümler eklem sayımlarını, radyografileri, laboratuvar testlerini, hasta anketlerini ve genel yöntemleri kapsamaktadır (Tablo 1). Bu yöntemler münferit RA vakalarını değerlendirmek ve izlemek amacıyla *American College of Rheumatology (ACR)* Temel Veri Seti (24), Hastalık Aktivite Skoru (DAS) (5;6) ve klinik hastalık aktivite indeksi (CDAI) (7) gibi çeşitli yöntemlerin (1) birleştirilmiş indeksleri haline dönüştürülmüştür. Yöntemler ve indeksler aşağıda ele alınacaktır.

a) Eklem sayımı: 1965 yılında geliştirilen ve elin metakarpofalangeal (MCP), proksimal interfalangeal (PIP) ve distal interfalangeal (DIP) eklemleri ile ayağın metatarsofalangeal (MTP) ve distal interfalangeal eklemlerini, omuz, dirsek, bilek, kalça, diz, ayak bileği, tarsal ve temporomandibüler, sternoklaviküler ve akromiyoqlaviküler eklemleri kapsayan 66/68 eklem sayımına dayalı pek çok eklem sayım yöntemi tarif edilmiştir (8). *American Rheumatism Association (ARA)* [artık *American College of Rheumatology (ACR)*] Terimler Sözlüğü Komitesi’ne göre eklemler beş değişken için 0, 1, 2, 3 olarak skorlanmıştır: Şişlik, hassasiyet, hareket esnasında ağrı, hareket kısıtlılığı ve deformite (9;10).

Eklem sayımının çeşitli uyarlamaları tarif edilmiştir. Ritchie İndeksi (11) hassasiyet açısından derecelendirilmiş 52 eklemi kapsamaktadır. 36 (12), 28 (13) ve 42 eklemi (14)

İLERİYE YÖNELİK VERİ TABANLARI VE KAYIT SİSTEMLERİ-NEDEN?

Theodore Pincus

*New York Üniversitesi,
New York, ABD*

içeren azaltılmış eklem sayımları, genellikle hastalıklı olmayan eklemlerin çıkarılması ile yoğun klinik ortamlarda ölçümü kolaylaştırmaktadır (13). Yeni tedavilere odaklı, çoğu eklem sayımı formatları inflamasyon göstergesi olarak yalnızca şişme ve hassasiyeti içermektedir. Hareket kısıtlılığı ya da deformiteyi dahil etmemeye yönelik uygulama uzun vadede eklem sayımının değerini kısıtlayabilir.

Bir gözlemci tarafından geleneksel hassas eklem sayımı ile önemli ölçüde ($\rho=0.44$ ila 0.87) ve şiş eklem sayımı ile daha alt düzeylerde uygunluk gösteren, hastanın kendisinin bildirdiği eklem sayımları bildirilmiştir. Romatolojide hastalık aktivitesinin hızlı değerlendirmesinden (RADAR) (15) uyarlanan bir hastanın kendisinin bildirdiği romatoid artrit hastalık aktivite indeksi (RADAI) (16) bir hekim/değerlendirici tarafından gerçekleştirilen hassas eklem sayımı ile önemli ölçüde uygunluk göstermektedir: ($\rho=0.55$, eritrosit sedimentasyon hızı (ESR) ile C reaktif protein (CRP) arasındaki korelasyondan daha yüksek bir korelasyon düzeyi (25)). Hastanın kendisinin bildirdiği eklem sayımları önümüzdeki yıllar içerisinde klinik araştırmalarda ve pratikte romatologlar tarafından belki de çok daha fazla kullanılacaktır.

Şiş ve hassas eklem sayımı RA hastalarında inflamatuvar aktiviteyi değerlendirmek için en özgün yöntemi sunmakta (26) ve klinik deneylerde aktif tedavileri kontrol tedavilerinden ayırt etmek için geleneksel anlamda en önemli yol olarak görülmektedir (27;28). Bir ACR 20/50/70 yanıtı hem şiş ve hem de hassas eklem sayımında % 20, % 50 ya da % 70 iyileşmeyi gerektirmektedir ve DAS hem şiş ve hem de hassas eklem sayımını içermektedir. Buna rağmen resmi eklem sayımlarında (Tablo 1) sıklıkla göz ardı edilen çok sayıda önemli kısıtlamalar görülmüştür.

Eklem sayımlarının yayımlanmış çalışmalarda gözlenen tekrarlanabilirliği çok düşüktür (29;33), ancak bu tekrarlanabilirlik çalışarak geliştirilebilir (32). Bir manyetik rezonans görüntüleme (MRI) taraması ya da ultrason, inflamasyonu ölçmek bakımından eklem muayenesine göre daha duyarlı olabilir (34). Etkilenen eklemlerin sayısı sıklıkla 5 ila 15 yıl içerisinde azalırken, çalışma maluliyetine ve

erken ölümlere neden olan eklem hasarı ve fonksiyonda gerileme görülmektedir (35;44); bu nedenle % 20 ve hatta % 50 düzeyinde bir iyileşme yine de zaman içerisinde ileri eklem hasarı ile ilişkili olabilmektedir (45;46).

Klinik deneylerde plasebo ya da kontrol tedavisi alan hastalarda iyileşme 7 ACR Temel Veri Setini oluşturan parametreler arasında şiş ve hassas eklem sayımları, genellikle hasta anketi skorları ve laboratuvar testleri gibi diğer RA Temel Veri Seti ölçümlerine göre daha hassastır (47). Eklem sayımlarının aktif ve kontrol tedavilerini ayırt etme becerileri hasta anket skorları ve laboratuvar testlerinden daha yüksek değildir (48;49). Resmi kantitatif eklem sayımları RA hastalarının çoğu vizitinde uygulanmamaktadır (50). Alışılmış romatolojik bakım açısından, muhtemel bir RADAI hastanın kendisinin bildirdiği eklem sayımı dahil olmak üzere, hastanın bildirdiği verilerini teyit etmek için kantitatif olmayan özenli bir eklem sayımı yeterli olabilmektedir.

b) Radyografiler: Radyografiler için ilk kantitatif skorlama sistemi 1949 yılında tarif edilen global IIV Steinbrocker skorudur (51). Bu global skor kolaylıkla hesaplanır, ancak zaman içerisindeki değişiklikleri görmek açısından yeterince ayrıntılandırılmamıştır. Bu nedenle Larsen (52;53) ve Sharp (54;55) tarafından RA'te el radyografilerini skorlamak için geliştirilen ve van der Heijde (56;58), Rau (59;60) ve diğerleri tarafından modifiye edilen metotlara dayanarak ayrıntılı kantitatif radyografik skorlar geliştirilmiştir. Tüm bu kantitatif skorlar birbirleri ile büyük ölçüde uyumludur. Van der Heijde tarafından modifiye edilen Sharp skoru klinik araştırmalarda günümüzde yaygın olarak kullanılmaktadır (58).

Radyografilerin RA'in durumunun değerlendirilmesinde bazı üstünlükleri vardır (61) (Tablo 1). Karakteristik erozyonların ya da simetrik eklem aralığı daralmalarının varlığı, RA için oldukça spesifikdir. Radyografi bunun yanı sıra zaman içerisinde iyi kontrol edilmeyen inflamasyondan kaynaklanan kümülatif hasarı da yansıtmaktadır. Radyografiler uzun dönemler boyunca karşılaştırma imkanı sağlayan "kalıcı" bir kayıt da oluşturabilmektedir.

Ancak, radyografilerin RA hastalarının

değerlendirilmesinde bazı önemli istenmeyen özellikleri de söz konusudur (Tablo 1). Klinik deneylerde 12 hafta süreyle aktif tedavi alan ve kontrol tedavisi alan büyük gruplar arasındaki farkları belirlemek mümkün olsa da, radyografler hastaların çoğunda yavaş değişmektedir ve değişiklikleri değerlendirmek için en az bir yıl ya da daha fazla bir süre gerekmektedir. Kantitatif radyografik skorların uygulanması sıkıcı bir iştir ve resmi klinik çalışmaların dışında nadiren uygulanmaktadır. Gerçekten de çoğu romatoloğun (ve radyoloğun) radyograflerin kantitatif değerlendirilmesinde deneyimi yoktur.

Radyografler eklem yıkımının önemli öngörü ve dökümantasyonunu sağlar. Ancak hasta anketlerindeki fonksiyonel durum ölçümleri, çalışma maluliyeti, maliyetler ve erken ölümler gibi ağır RA sonuçlarının kantitatif ön göstergeleri olarak radyograflere göre çok daha anlamlıdır (62;64). RA klinik araştırmalarında radyograflerin prognozunu iyi bir göstergesi olmaması böyle çalışmalarda radyografleri çekilen bölgelerin genellikle eller ve ayaklar olmasındandır. Halbuki dizler omuzlar ve kalçalar gibi büyük eklemler prognozda daha fazla etkili olabilirler (65). Uzun dönem RA sonuçları açısından büyük eklem radyografleri belki el ve ayak radyograflerine göre daha yüksek prognostik bir değer ortaya koyabilirlerdi. Yine de, pek çok romatolog radyografik yapısal değişikliği RA’te büyük prognostik öneme sahip “objektif” bir yöntem olarak kabul etse de, bildirilen kanıtlar radyografik verilerin RA sonuçlarının uzun dönem prognozunda oldukça az bir değere sahip olduğuna işaret etmektedir.

c) Romatizmal hastalıklarda laboratuvar testleri: 20. yüzyıl tıbbının baskın paradigması olan “biyomedikal model” (66) ile uyumlu olarak, romatizmal hastalıklarda laboratuvar testleri çoğu hekim ve hastanın laboratuvar verilerinin pek çok hastalığın tanı ve tedavisinde en kesin bilgileri sağladığına inanıyor olmasıdır. Romatoid faktör (67), antiskik sitrülün protein (anti-CCP) (68), eritrosit sedimentasyon hızı (ESR) (69), C reaktif protein CRP (70;71) testleri ve diğer testler dahil olmak üzere pek çok laboratuvar testi RA’li hastalarda yaygın olarak kullanılmaktadır. Anti-CCP ve romatoid faktör için pozitif bir serolojik test progresif RA tanısını doğrulamaya destek olabilmektedir (72).

>28 ESR değeri ya da anormal CRP sıklıkla klinik çalışmalara dahil edilme kriteridir (73). Kadında 30’dan az ve erkekte 20’den az ESR değeri ACR remisyon kriterlerinin karşılanması için gereklidir (74). ESR ve CRP’de azalmalar RA tedavilerinin tüm başarılı klinik deneylerinde hasta gruplarında görülmektedir ve kontrol tedavisi ile karşılaştırıldığında aktif tedavinin etkinliğini göstermektedir.

Buna rağmen romatizmal hastalıklarda laboratuvar testleri bazı önemli kısıtlamalar içermektedir (ki bunlar sıklıkla göz ardı edilmektedir) (Tablo 1). Öne çıkan bir kısıtlama sık görülen “yanlış pozitif” ve “yanlış negatif” sonuçlardır. Serum kolesterol, kreatinin, glikoz, hemoglobin A_{1c} vs. gibi hatırı sayılır ölçüde daha yüksek özgünlüğe sahip pek çok teste kıyasla hiçbir romatolojik kan testi hiçbir romatizmal hastalıkta % 100 pozitif ve tüm normal bireylerde % 100 negatif değildir. Örneğin farklılaşmamış erken artriti olan bir hastada bir anti-CCP testi RA’nin tanınmasında romatoid faktöre göre daha yüksek bir duyarlılığa sahipken (75), tedaviye ihtiyaç duyan RA’li hastaların üçte birinden fazlası (neredeyse % 40 kadarı) anti-CCP açısından negatiftir (romatoid faktör açısından da) (72). Negatif sonuçlar bazı hekimler tarafından RA olmadığı şeklinde hatalı yorumlanabilmektedir (72). Anti-CCP ve romatoid faktör patogenezi anlamak açısından ileri araştırmalar için büyük değer taşımaktadır, fakat klinik ortamda onlara verilen değer abartılıdır.

Aynı şekilde ESR ve CRP’nin RA’li hastaların yaklaşık % 40’ında normal olduğu önce Wolfe ve Michaud (76) tarafından bildirilmiş ve bu iki yeni çalışmada da doğrulanmıştır (77). Klinik iyileşme ile birlikte ESR veya CRP’de bir düşüş klinik deneylerde hasta gruplarında ve bazı münferit hastalarda görüldüyorsa da, ESR ve CRP klinik iyileşme olsa dahi pek çok hastada aynı düzeylerde kalabilmektedir (78).

Romatoid faktör, anti-CCP, ESR ve CRP için “yanlış negatif” testlerin yanı sıra romatizmal bir hastalığı olduğundan şüphelenilen pek çok bireyde “yanlış pozitif” testler de görülebilmektedir. Popülasyon araştırmalarında normal popülasyonun yaklaşık % 1’inde romatoid faktör pozitif bulunur. RA, popülasyonun yaklaşık

% 0.5'inde ortaya çıktığından, pozitif romatoid faktör testi, RA hastalığı olan insanlarda olduğu kadar olmayanlarda da aynı sıklıkta görülmektedir (72). Açıkçası romatoid faktör testleri herkes için istenmemektedir, ancak popülasyonda muskuloskeletal semptomların prevalansı % 15'tir (fibromiyalji % 5'tir). Gerçekten de popülasyon verilerine dayanarak romatoid faktör pozitif pek çok insanda muhtemelen RA olduğu kadar fibromiyalji de vardır.

Klinisyenler ve hastalar laboratuvar testlerine sıklıkla gereksiz bir önem verirler. Halbuki bu testlerin RA'de kullanımı önemli kısıtlamalara sahiptir. RA tanısının "altın standardı" herhangi bir laboratuvar testten ziyade klinik bir değerlendirmeye dayanır. Son olarak CDAI ile bir hasta hakkında karar verildiğinde hiçbir laboratuvar testine dayanmayız çünkü CDAI laboratuvar testi içermez (7).

d) Romatizmal hastalıklarda hasta anketleri:

Hastaların kendilerinin doldurduğu anketler RA'li hastaları değerlendirmek ve izlemek için son yirmi yıldır artarak kullanılagelmıştır. Sağlık değerlendirme anketi (HAQ) (79) 1980'de yayımlandığında romatolojide önemli bir kilometre taşı olmuştur. HAQ giyinmek, kalkmak, yemek yemek, yürümek, yıkanmak, uzanmak, kavramak ve günlük işleri yapmak gibi 20 yaşam aktivitesinin (GYA)'nin 8 kategoride incelenmesini kapsamaktadır. Hastaya 4 cevap şıkkı sunulur: "Herhangi bir zorluk olmaksızın" = 0, "biraz zorlukla" = 1, "büyük zorlukla" = 2 ve "yapamıyor" = 3. Sekiz kategorinin her biri için skor aynı kategorideki iki ya da üç GYA arasındaki en yüksek skordur; hastanın bu kategori için yardım ya da herhangi bir alet kullanması halinde skora 1 eklenmektedir (79). Toplam HAQ fonksiyonel maluliyet skoru her bir kategoriden alınan ve 0 ile 3 arasında değişen 8 skoru ortalamasıdır.

HAQ ayrıca ACR Temel Veri Seti'ndeki hasta tarafından bildirilen sonuç (HTBS) ölçümü olan ağrı ve global durumun değerlendirilmesi için 10 cm'lik görsel analog skalaları (VAS) da içermektedir (24;79).

HAQ RA'li hastaları değerlendirmek üzere

geliştirilmiş olduğundan "hastalığa özgü" bir anket olarak kabul edilmektedir. Buna karşın "jenerik" anketler, hastalığı olmayan insanlar da dahil olmak üzere her tür hastalığı olan hastalar üzerinde kullanılmak amacıyla geliştirilmiştir. En yaygın kullanılan "jenerik" hasta anketi pek çok hastalık türünde kullanılagelen kısa form 36'dır (SF36) (80). Jenerik anketler spesifik bir romatizmal hastalığın yaşam kalitesi üzerindeki etkilerini diğer romatizmal hastalıklar ve konjestif kalp yetmezliği ya da lenfoma gibi romatizmal olmayan kronik hastalıklar ile karşılaştırmak için kullanılabilmektedir.

Ne HAQ ne de SF36 bir klinisyen tarafından yoğun klinik ortamında hasta tedavi ederken gözden geçirilmek ve skorlanmak üzere tasarlanmamıştır. HAQ (79) standart bakımda klinik kararlar için hazır olması açısından kolayca skorlanabilen çok boyutlu HAQ (MDHAQ) olarak modifiye edilmiştir (81;82). Ayrıca MDHAQ'a dayanan ve hasta indeksi verilerinin rutin değerlendirmesi (RAPID) olarak adlandırılan ve fiziksel fonksiyon, ağrı ve global tahmin gibi 3 RA temel veri seti yöntemini içeren bir indeks geliştirilmiştir. Böyle bir indeks klinik deneylerde aktif ve kontrol tedavileri arasındaki farklılıkları DAS (83;85) kadar etkili bir biçimde ortaya çıkarmaktadır. RAPID3 10 saniyede skorlanabilmektedir (Yazıcı, 2008).

Fonksiyonel maluliyeti değerlendirmek amacıyla artrit etki ölçümü skalaları (AIMS) dahil olmak üzere ilave anketler geliştirilmiştir (5). Yaşam kalitesini değerlendirmek üzere kullanılan diğer anketler Avrupa yaşam kalitesi indeksini (EuroQoL) de içermektedir (86). Ağrının (87;88) ve yorgunluğun (89) değerlendirilmesinde basit VAS'nin ötesinde daha ayrıntılı anketler kullanılabilmektedir. "Her bir hasta için kişiye özel" olmak üzere McMasterToronto artrit hasta tercihi anketi (MACTAR) indeksi ya da hasta değerlendirme tekniği (PET) indeksi gibi (90;91) bireyselleştirilmiş yöntemler de vardır (90;91). RA dışındaki romatizmal hastalıklar için "hastalığa özgü" Western Ontario McMaster osteoartrit indeksi (WOMAC) (92) ve fibromiyalji etkisi anketleri de vardır (FIQ) (93). Son zamanlarda, fazla soruları azaltmak amacıyla bir soruya yanıtın uygun ilave soruları tetiklediği ve fiziksel fonksiyon için farklı maddelere dayalı standart bir skor ortaya

koyan “bilgisayar destekli teknoloji”yi (CAT) kullanan bireyselleştirilmiş anketler de geliştirilmiştir (94).

HAQ’a karşın MDHAQ ya da SF36’ya karşın kısaltılmış bir sürüm olan SF12 örneklerinde olduğu gibi standart bakım için tasarlanmış hasta anketleri araştırma için tasarlanmış hasta anketlerinden oldukça farklıdır (95). Araştırma anketleri genellikle uzundur ve değerlendirmeye uygun değildir ya da klinik bakım esnasında kolaylıkla skorlanamamaktadır. Bu nedenle anket hasta ve klinisyene ilave bir yük getirirken, bakıma katkı sağlamadığı görülmektedir.

Standart bakım için tasarlanmış hasta anketleri aşağıdaki özelliklere dayanarak klinisyene zaman kazandırmalı ve hasta vizitlerinin kalitesini iyileştirmelidir (96;99):

- Hastaların çoğu tarafından 5 ila 10 dakika içerisinde tamamlanmalıdır;
- Bir klinisyen tarafından 5 ila 10 saniye içerisinde değerlendirilebilmelidir;
- Skorlamayı kolaylaştıracak şekilde tasarlanmalıdır, genellikle anket üzerinde skorlama şablonları da bulunmalıdır;
- 10 ila 20 saniye içerisinde skorlanmalı ve bir şemaya girilmek üzere hazır hale gelmelidir;
- Her türlü romatizmal hastalığı olan hastalar için aydınlatıcı olmalıdır;
- Tıbbi öyküye ait veriler de (sistemlerin analizi gibi) içerebilir.

Neredeyse işin tamamı hekim, personel ya da hastalar tarafından yapılmalı, hekim yalnızca verileri gözden geçirmek için çok az bir zaman harcamalıdır (96;100).

Yukarıda da belirtildiği gibi, hasta anket skorlarının aktif ve kontrol tedavilerini ayırt etme becerileri genellikle şiş ya da hassas eklem sayımlarıyla benzer ya da daha fazladır (48;49). Bu nedenle bir kantitatif hasta anket skorları indeksi klinik çalışmalarda DAS ya da ACR kriterleri kadar etkilidir (83;84;101). Ayrıca hasta anketleri – eklem sayımları, radyografik skorlar ya da laboratuvar testleri değil – fonksiyonel durum (36;102), çalışma maluliyeti (103;105), maliyetler (106), eklem replasman ameliyatı (107) ve erken ölüm (36;65;108;113) dahil olmak üzere RA’li hastalardaki tüm ağır uzun dönem sonuçları açısından en önemli öngörülerini içerirler.

Klinik romatolojideki çoğu karar laboratuvar ya da görüntüleme verilerinden ziyade klinik olaylar temelinde verilir. Bu nedenle hastaların kendilerinin verdiği bilginin klinik tedaviyi yönlendirmek açısından, laboratuvar ya da görüntüleme çalışmalarının “objektif” yöntemlerine göre daha değerli olması şaşırtıcı değildir. Hasta anketleri başlangıç skorları sağlamak, değişikliği izlemek ve iyileşmeyi ve tedavinin değerini belgelemek amacıyla “kantitatif öykü” sağlar.

Hasta anketlerinin eksik yönleri de vardır. Çoğu anket alışılmış klinik bakım için fazla hantaldır ve yalnızca araştırmalarda kullanılmaktadır. Böyle anketlerin kısaltılmış sürümleri tasarlanmıştır (28;114). RA’te şiş eklem sayımları, polimiyozitte kas enzimleri, SLE’de antiDNA antikörlerine kıyasla hasta anketleri inflamatuvar aktivitenin derecesini değerlendirmeye uygun değildir. Hasta anketleri kültürel ve dil farklılıklarıyla sınırlı olabilmektedir, örneğin İspanyol kökenli hastaların ağrı skorları beyaz hastalara göre daha yükseken, Asya kökenli hastalar beyaz hastalara göre daha düşük ağrı skorlarına sahiptir (115). Hasta anketlerinin belirli yanıtların arzu edilen bir sonucu doğurmasının muhtemel olduğunu düşünen hastalarca saptırabileceği de düşünülse bile bu tür bir teorik sorun klinik bakımda pek önemli değildir. Hekimlerin, hedeflenen sonuçlara yönelik olarak değerlendirme yapmaları daha olası görülmektedir. Ancak bu sorunun gelecekte daha fazla problem yaratacağını düşünenler vardır.

e) Hasta ve hekim tarafından global durum değerlendirmesi: Global durum hasta ya da hekimin tahminlerine göre değerlendirilebilir. Geçmişte global durum *Steinbrocker American Rheumatism Association* (artık *American College of Rheumatology*) fonksiyonel sınıfı gibi bir 4 nokta skalası üzerinden ölçülmekte idi (51). Ancak hastaların % 70’i Sınıf II iken ve yıllar içerisinde anlamlı iyileşme ya da ilerlemeye rağmen yine Sınıf II’de kaldıklarından, böyle bir 4 nokta skalası değişime oldukça duyarlıdır.

Daha yeni olarak hasta ya da hekim /değerlendirici tarafından bildirilen global durum 10 cm VAS’a göre ölçülmektedir. Hastanın kendisinin bildirdiği global durum değerlendirmesi

için HAQ üzerinde 10 cm'lik çizgi şeklinde ve MDHAQ'ın güncel sürümleri üzerinde 21 halka şeklinde bir VAS konmuştur. 21 halkadan oluşan global bir VAS geleneksel 10 cm'lik çizgi VAS'ı kadar işlevseldir (116) (ağrı ve yorgunluğun değerlendirilmesi açısından da).

Global durumu 0–10 skalası üzerinden değerlendiren skalalar 4-nokta skalalara göre değişkenliğe oldukça hassastır. RA'ın değerlendirilmesi için ACR temel veri seti hem hastanın global durum tahmini için bir 10 cm VAS, hem de hekimin/değerlendiricinin global durum tahmini için benzer ikinci bir VAS içermektedir. Bu global tahminler tüm klinik çalışmalar da vardır. Klinik çalışmalarda hekimin global durum tahmininin aktif tedaviyi kontrol tedavisinden ayırt etmede şiş eklem sayısı ve hassas eklem sayımından, hastanın global durum tahmininin ise daha ayrıntılı olan HAQ skorundan (ya da VAS ağrı skorundan) daha büyük bir etkinliği olduğu gözlenmiştir (48).

Global VAS'ın başlıca sınırlayıcılığı spesifikliktir – global durumda iyileşme ya da kötüleşme, RA durumunda şiş ve hassas eklemler anlamındaki doğrudan iyileşmeyi her zaman yansıtmayabilir. Bu sınırlayıcılık RA'deki çoğu değerlendirme yöntemi için geçerlidir, örneğin ağrı skoru kırık ya da yaralanma ile artmış olabilir ya da ESR inflamatuvar aktivitedeki bir artıştan ziyade enfeksiyon ya da lenfoma nedeniyle artmış olabilir. Bu gözlemler klinik kararları verirken hekimlere ya da sağlık profesyonellerine duyulan ihtiyacı vurgulamaktadır.

f) Romatizmal hastalıkları olan hastaları değerlendirmek için kullanılan hastalık aktivitesi ve hasar indeksleri: Yukarıda da belirtildiği üzere, romatizmal hastalıklarda bir “altın standart” yönteminin yokluğu, RA için 3-7 yöntemi birleştiren indekslerin geliştirilmesine yol açmıştır. Bunlar 7 yöntem içeren ACR Temel veri setine (2-4) dayanmaktadır. ACR temel veri seti 3 adet veri değerlendiriciden – hassas eklem sayısı, şiş eklem sayısı ve hekim/değerlendirici global durum tahmini; 1 laboratuvar testi ESR ya da CRP; ve 3 hastanın kendisinin bildirdiği veri – fiziksel fonksiyon, ağrı ve hasta global durum tahminini kapsar. DAS ise bu yöntemlerin 4'ünü içermektedir –

hassas eklem sayısı, şiş eklem sayısı, laboratuvar testi ve hasta global durum değerlendirmesi. Skorlar DMARD tedavisinin değiştirildiği hasta ziyaretlerinin ayrıştırıcı analizlerinden geliştirilen karmaşık matematiksel bir formülünden türetilmiştir ve skorlama bir DAS hesaplayıcısı ya da bilgisayar web sayfası gerektirmektedir (5;6). CDAI hassas eklem sayısı, şiş eklem sayısı, hekim /değerlendirici global tahmini ve hasta global tahminini kapsamaktadır (7). Skorlama toplam 0-76'lık olmak üzere hassas ve şiş eklemlerin bir 0-28 arasında sayısı ve 0-10 arasındaki hasta ve hekim/değerlendirici global VAS skorları bazında yapılıp ve hesaplayıcı ya da bilgisayar gerektirmemektedir.

Yalnız hasta tarafından verilen bilgilere (HTBS yöntemi) dayalı olarak da indeksler geliştirilmektedir (85;117). ACR temel veri setindeki 3 HTBS yönteminin indeksi – fiziksel fonksiyon, ağrı ve hasta global durumu – hasta indeks verilerinin rutin değerlendirilmesi (RAPID3) olarak anılmaktadır. RAPID3 ACR veya DAS28 kriterlerine benzer düzeylerde leflunomid (83;84), metotreksat, (83;84), adalimumab (101) ve abatacept (118) içeren aktif tedaviler ile kontrol tedavilerini ayırt etmektedir ve bu deneylerdeki ve hasta bakımındaki DAS 28 ile önemli ölçüde uygunluk göstermektedir (85). RAPID3'ün uygulanması aşağıda daha ayrıntılı olarak tarif edilmektedir (Bölüm 3).

B. RANDOMİZE ÇALIŞMALARIN EKSİK YÖNLERİ

Randomize çalışmalarda çeşitli gözlemciler tarafından (119-130) anlatılmış bazı eksik yönlere dikkat çekmek istiyorum. Bu kısıtlamalar araştırmacılar, sağlık profesyonelleri ve toplum tarafından sıklıkla göz ardı edilmektedir. Elbette randomize klinik çalışma herhangi bir tıbbi girişimin sonuçlarını başka bir müdahale ya da bir plasebo ile karşılaştırmak için en uygun yaklaşımdır. Randomize kontrollü klinik çalışma tek bir değişkenin; çalışmada kullanılan ilaç tedavisinin, izolasyonuna izin vermektedir ve bu da “bilimsel çalışma” yöntemine uyar.

Randomize kontrollü klinik çalışmanın üstünlüğü, “yeni bir tedavinin değerini değerlendirmek için yegane yaklaşım” olduğu şeklinde bir ifade ile tıp literatüründe sıklıkla

dile getirilmektedir. Çalışmalardan elde edilen veriler “1. derece” kanıt olarak kabul edilebilmekteyken, gözleme dayalı araştırmalar genellikle “3. derece” kanıt olarak kabul edilmektedir. Ancak randomize kontrollü klinik çalışmalarda özellikle de uzun süren kronik hastalıkları ilgilendirdiklerinde bir dizi önemli sorun görülmektedir.

Benim eleştirilerimin çoğu 35’in üzerinde klinik çalışma yürütme tecrübeme dayanmaktadır ve randomize kontrollü klinik çalışmalara kesinlikle inanmaktayım. Yine de randomize kontrollü klinik çalışmaların kısıtlamaları longitudinal çalışmalar için bir başka mantıksal temel sağlayabilir. Bu kısıtlamalar iki kategori altında toplanabilir, modern klinik araştırmada pragmatik (işlevsel) uygulamadan kaynaklananlar ve klinik çalışma metodolojisinin doğası gereği olanlar.

a) Randomize kontrollü klinik çalışmaların sekiz işlevsel eksikliği: Klinik çalışmaların sekiz eksikliği “işlevsel” sınırlama olarak kabul edilmektedir. Zira çağdaş klinik çalışmaların tasarımı pratik düşüncelerden kaynaklanmaktadır. Teoride ideal bir klinik çalışma ile bu eksiklerin üstesinden gelinebilir, ancak böyle bir ideal klinik çalışma bugüne kadar henüz tasarlanmamıştır.

1) Dışlama kriterleri bir klinik çalışma için uygun olduğu düşünülen, belirli bir tanı almış hastaların çalışmaya alınabilmesini % 10’dan daha azını çalışmaya dahil edecek şekilde sınırlayabilir.

Bir klinikte aynı tanıyı almış tüm bireyler ve hastalar teorik olarak bir klinik çalışmaya katılmaya uygun olabilir. Ancak hastanın yaşı, eğitim düzeyi, hastalığın ağırlığı, komorbid hastalıkların varlığı, önceki ve eşlik eden müdahaleler ve daha pek çok değişken olası sonuçları etkileyebilmektedir. Bu nedenle tüm klinik çalışmalar nispeten homojen bir hasta grubunu araştırmak adına “dahil edilme” ve “dışlanma” kriterlerini tanımlar.

RA’te dahil edilme kriterleri hastaların öncelikle, 6’dan fazla şiş eklem, 6’dan fazla hassas eklem, 28’den yüksek ESR gibi iyileşmeyi fark etmeye yetecek kadar ağır görünen hastalık aktivitesine sahip olduklarını sağlamak üzere

tasarlanmıştır. Bu uygulama RA klinik çalışmalarına katılabilmeyi yıllarca kısıtlamıştır. Hatta RA’li hastaların durumu 1985 ile 2000 yılları arasında belgelendiği üzere (137-140) anlamlı bir şekilde iyileştiğinden, son yıllarda bu sorun daha da belirgin hale gelmiştir. Örneğin 2001 yılında kliniğimdeki bakılan hastaların % 5’den azı ATTRACT çalışmasına (141) dahil olabilmenin basit kriterlerini doldurmaktaydı. Ayrıca aynı şehirde özel bir muayenehanede görülen RA erken dönem hastaları arasında % 20’den azı erken romatoid artrit (ERA) çalışmasına katılmaya uygun bulunmuştur (73).

Bu gözlemler ilaç şirketlerini çoğunlukla Doğu Avrupa ya da Güney Amerika gibi RA tedavisinin daha az geliştiği bölgelerde daha yeni çalışmalar yürütmeye yöneltmiştir. Fakir ülkelerdeki hastalar, zengin ülkelere göre daha kötü klinik durumdadır. Gerçekten de farklı ülkelerin milli geliri ile bu ülkelerdeki RA’li hastaların klinik durumu arasında güçlü bir korelasyon söz konusudur, bu durum 20 ülkede 60 alanın her birinden 100 hastayı içeren uluslararası bir veri tabanı olan Romatoid Artritli Hastaların Standart Klinik Bakım (QUEST-RA) Çalışmasının Kantitatif Hasta Anketleri’nde de belgelenmiştir (142).

Dışlama kriterleri klinik çalışmalara katılımı, ağır komorbiditesi olan hastaların çalışmaya alınmaması nedeniyle de kısıtlamaktadır. Randomizasyon süreci bir klinik çalışmanın sonuçlarını etkileyebilecek diğer değişkenlerin istatistiksel ayarlamasına izin verirken, abartılı dahil etme ve dışlama kriterleri, çoğu RA klinik çalışmasının sonuçlarının alışılmış hastaların çoğuna genelleştirilmesini zorlaştırmaktadır (136).

2) Kronik hastalıklarda klinik çalışmaların oldukça kısa sürmesi

Kronik hastalıklarda klinik çalışmaların en önemli kısıtlamalarından biri, yalnızca uzun sürelerde gelişen anlamlı klinik eğilimleri fark etmek için yetersiz olabilen kısa gözlem süresidir. Bu konuda çarpıcı bir örnek sistemik lupus eritematozus (SLE) nefritini önlemek için prednizon artı siklofosfamide karşı yalnızca prednizonun karşılaştırılması amacıyla yürütülen bir çalışmada görülmektedir. Prednizon artı siklofosfamidin avantajları 10 yılı aşkın süre ile

izlenmiş ve yalnızca prednizon ile tedavi edilen hastaların yaklaşık % 30'una karşın siklofosfamid + prednizon alan hastaların yaklaşık % 90'ında böbrek fonksiyonlarının korunduğu görülmüştür (143). Ancak bu çalışmada renal fonksiyon gözlemin ilk dört yılında her iki grupta da hastaların % 90'ından fazlasında korunmuştur ve gruplar arasındaki farklılıklar ancak dört yıl sonra ortaya çıkmıştır. Şayet bu çalışma romatoloji ile ilgili yayımlanmış diğer tüm klinik çalışmalarda olduğu gibi, yalnızca üç yıl ya da daha kısa bir süre sürdürülmüş olsaydı, çalışma sonucu prednizon artı siklofosfamid tedavisinin yalnızca prednizon ile tedaviye göre bir avantajının olmadığı şeklinde olacaktı! Çoğu klinik ya da hastanede yürütülemeyecek olan bu klinik çalışma için yalnızca *National Institutes of Health (NIH)* gibi bir yerel kurum kaynak sağlamıştır.

RA klinik çalışmalarında kısa sürenin ve hasta seçiminin sınırlayıcılığı, 1990'da bildirilen ve sulfasalazin, D-penisilamin, metotreksat ve enjekte edilebilir altının birbirlerine karşı benzer etkinliğinin olduğunu gösteren klinik çalışmaların bir metaanalizi ile açıklanabilir (144). 538 hasta üzerinde 1083 DMARD kürünün uzun dönem sonuçlarının analizi, beş yıl sonra 56 azatiyoprin kürünün, 269 enjekte edilebilir altın tuzu kürünün, 193 penisilamin kürünün ve 228 hidroksiklorokin kürünün % 20'sine karşın 253 metotreksat kürünün yaklaşık % 70'inin sürdürüldüğünü göstermiştir (145). Çoğu klinik çalışma ile alışılmış tedavi arasındaki iki önemli fark, klinik çalışmaların genellikle önceden DMARD tedavisi almayan hastalar üzerinde ve yalnızca bir yıl boyunca yapılması ve günlük tedavi analizinin ise önceden herhangi bir DMARD tedavisi alan tüm hastalar üzerinde ve beş yıl boyunca sürdürülmüş olmasıdır. Klinik çalışmaların koşulları alışılmış klinik tedavi kürlerinin alt gruplarına uygulandığında, (ilk DMARD bir yıl boyunca), DMARD kürlerinin sürdürülmesi arasında, klinik çalışmaların sonuçlarına benzer şekilde, herhangi bir fark görülmemiştir. Bu örnek çoğu çalışmanın hasta seçimi ve kısa süreden kaynaklanan kısıtlamalarını ortaya koymaktadır.

3) İstatistiksel güç açısından yetersiz hasta sayısı nedeniyle, klinik olarak önemli farklılıklar istatistiksel olarak anlamlı çıkarmaz

Pek çok romatizmal hastalık, 2000 kişide 1'den az görülen "ender" hastalıklar olarak sınıflandırılmaktadır. Makul bir süre için bir klinik çalışmada yer alacak yeterli sayıda hasta bulmak zor olabilir. Bu nedenle klinik çalışmalar tedavi etkisi gösteremeyebilirler zira sadece yetersiz bir istatistiksel güç söz konusudur. İstatistiksel güç birden fazla klinik çalışmanın metaanalizi ile artırılabilir.

Örneğin SLE nefriti için immünsüpresif ilaçlar artı prednizona karşı yalnızca prednizonu karşılaştırmak amacıyla 8 küçük klinik çalışma gerçekleştirilmiştir. Bu çalışmaların çoğu immünsüpresif ilaçlar artı prednizonun üstünlüğünü salt prednizona karşı istatistiksel olarak anlamlı bulmamıştır (146). Ancak bir metaanalizde yalnızca prednizon ile karşılaştırıldığında, immünsüpresif ilaçlar artı prednizonun istatistiksel olarak anlamlı bir avantajı görülmüştür (146). Bu çalışmaların yalnızca bir ya da ikisi mevcut olsaydı, sonuç prednizon+immünsüpresiflerin herhangi bir avantaj sağlamadığı şeklinde olabilirdi. Yine de metaanalizin olası tuzakları unutulmamalıdır.

4) İstatistiksel anlam mutlaka klinik anlam ya da önemi göstermemektedir

Hasta sayısının azlığı klinik olarak anlamlı bir farkın istatistiksel olarak anlamsız görünmesine neden olabileceği gibi çalışmanın çok sayıda hastada yapılması klinik olarak anlamsız bir farkın istatistiksel olarak anlamlı görünmesine neden olabilir. Bu kısıtlama 1970'lerde ve 1980'lerin başlarında yürütülen ve ibuprofen, naproksen, piroksikam ve diklofenak gibi çeşitli non steroid anti-inflamatuvar ilaçların (NSAID) RA'li hastalarda hassas veya şiş eklem sayısı bakımından bir iyileşmeye yol açtığı yüzlerce klinik çalışmada görülmüştür. Veriler çok sayıda hastada istatistiksel olarak çok anlamlı iken, bu ajanlar çoğu hastaya marjinal klinik fayda sağlamıştır (Goetzsche, 1990 301 /id).

5) Hastanın bir ilaç çalışmasında hangi gruba randomize olduğu dışında önemli değişkenler klinik çalışma makalelerinde genellikle göz ardı edilmektedir.

Randomize kontrollü klinik çalışmanın temel tasarımı bir müdahaleyi bir diğerine ya da plaseboya karşı kullanarak sonuçlardaki

farklılıkları tanımlamaya odaklanmıştır ve sonuçların bildirilmesi doğal olarak bu karşılaştırmayı vurgular. Ancak bazı çalışmalarda sonuçlar hastanın hangi ilaca veya plaseboya mı randomize edildiğinin dışında değişkenler tarafından daha fazla etkilenmektedir. Bu fenomene iyi bir örnek, henüz yeni kalp krizi geçirmiş hastalarda ikinci bir kalp krizini önlemek amacıyla bir beta bloker ilaca (propranolol) karşı plasebo tedavisini karşılaştırmak üzere tasarlanmış Beta Bloker Kalp Krizi (BHAT) çalışmasıdır. Çalışma propranololün plasebodan daha etkili olduğunu gösterirken, hastanın temel eğitim düzeyi, kişisel bakım için yardımcısı olması, yaşam stresi ve sosyal desteği, hastanın ilaç ya da plasebo almış olsun ya da olması mortalitede daha büyük farklılıklar ile ilişkili bulunmuştur (147;148). Ancak eğitim düzeyine göre farklılıklar ilacın kendisine göre farklılıklar kadar iyi bilinmez.

6) Esnek olmayan dozaj ve çekilme kriterleri

Çoğu randomize kontrollü klinik çalışma protokolleri, mümkün olduğunca tutarlı şartlar oluşturma arzusuna bağlı olarak esnek olmayan bir dozaj programı ve çalışmadan çekilmek için esnek olmayan koşullar içermektedir. Örneğin erken metotreksat klinik çalışmaları düşük dozlar ve karaciğer fonksiyon testlerinde hafif artışlar gösteren hastaların çalışmadan çekilmesi koşulunu içermektedir. Bu protokol koşulları, alışılmış tedavide uygulandığında metotreksatın etkinliğine olması gerekenin altında ve advers olaylarına olması gerekenin üzerinde değer biçilmesine neden olmuştur (149;150). Gerçek etkinlik ve advers olayların çarpıtılması, gerçek klinik uygulamada ayarlamaları yansıtmayan esnek olmayan dozajdan kaynaklanıyor olabilir.

7) Hastalık sonuçlarının optimalin altında belirteçleri olabilen vekil belirteç (surrogate marker) gereksinimi

RA'nın ağır sonuçları eklem harabiyeti, eklem protezi gereksinimi, maluliyet ve erken ölümü kapsamaktadır ve hastalarda yıllar sonra ortaya çıkmaktadır. Bu nedenle 6-24 aylık klinik çalışmalarda kullanılan ölçümler uzun dönemli hasara yol açtığı düşünülen inflamasyonun kısa dönemli vekil belirteçleridir. Her ne kadar belirteçler uzun dönem ağır sonuçların

önlenmesi ile anlamlı bir şekilde uyumlu ise de, bu korelasyon çoğu zaman mükemmelin altında ve bazı olgularda da anlamlı değildir. Örneğin ağrı ve eklem hassasiyeti, RA temel veri seti içinde klinik çalışmalardaki değişikliklere duyarlı belirteçlerdir. Radyografik eklem hasarı da böyle bir belirteçtir ancak ilk 2 belirteçle (ağrı ve eklem hassasiyeti) radyografik değişimin iyi bir korelasyonu yoktur. (13). Ayrıca eklem hassasiyetinin doğal öyküsü beş yıllık bir dönem içerisinde iyileşmek iken, eklem harabiyeti ve buna bağlı deformite ve kısıtlı fonksiyonel kapasite artabilmektedir (111). Bu nedenle bir klinik çalışmada belirteçlerin gelecekteki hasarı ön görme kapasiteleri kısıtlı olabilir.

8) Çoğu çalışmada nadir advers olaylar tanımlanamamaktadır

Klinik çalışmaların tasarımı ve harcamaları çalışmaya katılabileceklerin çok küçük bir azınlığının klinik çalışmaya katılmasına elvermektedir; bir ilaç ile herhangi bir zamanda tedavi görmüş hastaların % 0.001'inden azı. Temsili bir örneklem ile dahi aşılamayacak özel bir kısıtlama ise ender advers olaylardır. Örneğin ağır bir yan etki 10000 hastada 1 meydana geliyor ise ve bir ilaç ruhsat almasından önce yalnızca 5000 hasta üzerinde çalışıldı ise, bu olası advers olayın çalışmalarda gözlenmeyecek olma olasılığı söz konusudur. Bu nedenle belirli bir ilacı alışılmadık ancak ağır advers olaylar açısından izlemek amacıyla en az 100000 hastaya dair gözlem verileri toplamak gereklidir.

b) İç kısıtlamalar: Yukarıdaki kısıtlamalar, bir hastalığı olan tüm kişilerin çalışmaya dahil edilmesi, esnek dozaj programı ve çok büyük denek sayıları gibi "ideal" şartlar altında aşılabılır. Kontrollü klinik çalışmaların bazı kısıtlamaları ise metodolojiden kaynaklanır gibi görünmektedir. Bu kısıtlamaların ideal klinik çalışma şartlarında dahi aşılması olası görünmemektedir.

1) Kontrol grubu olmasına rağmen, bir klinik çalışmanın tasarımı sonuçları büyük ölçüde etkileyebilmektedir

Klinik çalışmanın tanımlayıcı özelliği redüksiyonizm temelinde, bir laboratuvar deneyiminde olduğu gibi tek değişkene odaklamak için bir girişim grubu ile karşılaştırılmak üzere

bir kontrol grubu izlenmesidir. Bu tasarım, alışılmış klinik tedavinin daha ağır hasta olan kişilerin yeni tedaviyi almalarının, yol açtığı yanlılığı (bias) aşmaktadır. Gerçekten de yanlılık hastaların randomize kontrollü bir çalışmaya alınmasının ardından ortadan kalkar.

Ancak çalışma protokolünün kendisi, sonuçları çalışma tedavisi ile ilgili belirli bir çıkarımın lehine ya da aleyhine güçlü bir şekilde etkileyebilmektedir. Örneğin RA'te yeni bir biyolojik tedaviyi plasebo ile karşılaştırmak üzere, daha önce herhangi bir tedavi görmemiş veya daha önce iki standart tedavide "başarısız" olmuş hastalar üzerinde oynayacak, basit bir klinik çalışma düşünün. Klinik araştırmalar (ve sağduyu) genel olarak bir ilacın başka ilaçlara iyi yanıt vermemiş olan bir hastada üçüncü ilaç olarak kullanılmasındansa, aynı hastada ilk ilaç olarak kullanıldığı takdirde, etkinlik göstermesinin tanım üzere olası olduğunu göstermiştir. Bir ilacın üçüncü tedavi olarak denenmesi istendiğinde daha dirençli hastalar seçilmektedir. Yeni bir ilaç, bir hastalık için ilk tedavi olarak kullanılıyor ise etkinlik açısından plaseboya karşı istatistiksel olarak daha anlamlı farklılıklar gösterebilir fakat ilacı üçüncü tedavileri olarak alan hastalarda yalnızca marjinal (ve istatistiksel olarak anlamsız) olarak görülebilir.

Bir klinik çalışmanın tasarımı, sonuçları önceden belirleyemez. Buna rağmen bir klinik çalışmanın tasarımı bir müdahalenin plaseboya göre daha etkin olup olamayacağı olasılığını büyük ölçüde "olumsuz etkileyebilir".

Önemli nokta, sadece "kontrol" grubu bulunmasının randomize kontrollü klinik çalışmalarda, her çalışmanın içerisinde olan yanlılığı ortadan kaldıramamasıdır.

2) Klinik çalışmalar verileri gruplar halinde bildirmekte ve genellikle hastadan hastaya olan değişkenliği gözardı etmektedir

Kesin bir sonuç arzu edilse de, klinik çalışmalarda bir tedavi ya da diğeri lehine hatta plaseboya kıyasla bir tedavi lehine net bir karar çıkması sık görülmez. Örneğin kalça ya da diz osteoartritli hastalarda asetaminofene karşı diklofenak sodyum/misoprostol (151) ya da selekoksib (152) ile yapılan çapraz klinik

çalışmalarda bireylerin % 60'ı diklofenak sodyum/misoprostol ya da selekoksibi tercih etmiş ve % 20'si asetaminofen tercih etmiş iken, % 20'sinin diklofenak/misoprostole karşı asetaminofen ya da selekoksibe karşı asetaminofen arasında herhangi bir tercihi olmamıştır. Doğru bir yorum, diklofenak/misoprostolün kalça ya da diz osteoartritli çoğu hasta için daha yüksek etkinliğe sahip olduğu fakat bazı hastaların asetaminofen ile daha yüksek bir etkinlik gördükleri şeklinde olmalıdır. Ancak çoğu kez yorum tercih edilen ilacın tüm hastalar için optimal olduğu yönünde yapılır.

İnfeksiyonlara etkili antibiyotikleri sınamak için prototip klinik çalışmalar gerçekleştirilmiştir. Bakteriler tek hücreli basit organizmalardır ve tüm hastalar için spesifik bir bakteriyi ortadan kaldıracak tek bir optimal ilaç tanımlanmış olabilir. Buna karşın çok hücreli insan organlarında gastrik asidin azaltılması, kan basıncının kontrol edilmesi, ağrı ile baş edilebilmesi, depresyonun azaltılması vs. için tasarlanmış tedavilerde bir ilaca yanıtta varyasyon beklenmektedir. Bu koşullar altında farklı bireylerin farklı tedavilere farklı yanıt vermesi beklenirken, genellikle klinik çalışmaların belirli bir insan için en etkili olabilecek tedaviyi belirlemekten ziyade tüm hastalar için "en iyi" tedaviyi belirleyeceği görüşü vardır. Bir birey için optimal tedaviyi belirlemek açısından genom çalışmalarıyla "kişiselleştirilmiş tıp" bu yaklaşımı değiştirebilir. Ancak günümüzde belirli bir rahatsızlığı olan tüm hastalar için tek bir "en iyi" tedavi söz konusu değildir. "Kafa kafaya" çalışmalarda (Head-to-Head: Klinikte aynı tedavi alanında kullanılan standart tedavilerin karşılaştırması) bile hastaların çoğu tarafından tercih edilmeyen bir tedaviye yanıt veren bazı hastalar vardır.

3) Advers olayların yorumlanması standardize edilmemiştir ve büyük ölçüde farklılık gösteren kişisel yoruma bağlıdır

Tüm ilaçlar basit kaşıntıdan böbrek hasarına kadar uzanan belirli bir advers etkiyle ilişkili olabilir. İki ilacın karşılaştırılmasında A ilacının hastaların % 90'ında çok az yan etki ile remisyona neden olduğu fakat 10000 hastada 1 hastada böbrek yetmezliği yaşandığını varsayalım. Buna karşın B ilacı hastaların % 50'sinde herhangi bir

remisyon olmaksızın fakat herhangi bir ciddi advers olay olmaksızın iyileşmeye neden olmaktadır. Hangisi tercih edilen ilaçtır? Bu, risk ve faydaların nasıl değerlendirildiğine bağlıdır ve hiçbir komite bunu her bir birey için karara bağlayamaz.

Genel olarak hekimler sonuçları yorumlamak için bir klinik çalışmayı bildiren araştırmacıların kararına güvenir. Ancak gerçek uygulamada bazı hastalar bambaşka bir alternatif tercih edebilir ve optimal tedavi açısından “aydınlatılmış bir karar” verebilir. Bir klinik çalışma bir “siyah ve beyaz” tercih ima ederken, gerçek sonuçlar grinin tonlarını öne sürmektedir. Bu, yoruma bağlı duruma eşlik eden önyargılar ile birlikte, klinik çalışma verilerinin yorumlanması ve belirli yeni ilaçların ruhsatlandırılmasında A.B.D. Gıda ve İlaç Dairesi’nin (FDA) danışma kurullarının üyeleri arasında zaman zaman yaşanan tartışmaları da etkileyebilir.

4) Bir klinik çalışmanın metodolojisi hastaların “en iyi” tedaviyi alamayabilecekleri konusunda bilgilendirilmemeleri nedeni ile “plasebo etkisi”ni tehlikeye atmaktadır

Son on yılda herhangi bir hasta müdahalesinde “plasebo” etkisi ile ilgili hatırı sayılır veriler toplanmıştır (153-157). 20. yüzyıla kadar hekimlerin elinde çok az sayıda etkili ilaç vardı ve yine de sağlık profesyonellerinin pek çok durumda “tedavi edici” ilaç verdiklerine inanılıyordu. Bu plasebo etkisi, açıktır ki bir hekim bir hastayı en iyi tedaviyi saptamak amacıyla hastanın tedavi alabileceği ya da kontrol olarak hizmet edebileceği bir bilimsel çalışmada yer almaya davet ettiğinde tehlikeye girmektedir. Klinik çalışmaların savunucuları hastanın en iyi tedaviyi almasına olanak vermenin en etik yol olacağını ifade etmektedirler. Bu görüşe katılıyorum; buna rağmen hasta açısından bakıldığında, daha zayıf sonuçların elde edildiği gruba tayin edilme ihtimali de bulunmaktadır. Hatta hastaların bir kontrol grubuna dahil edilebilecekleri bilgisini vermenin negatif etkileri olabileceği de ifade edilmiştir.

Klinik çalışmaların çoğu, her şeye rağmen klinik çalışma metodolojisi içerisinde “plasebo etkisi”nin önemli olduğundan yola çıkarak, gerek plasebo, gerekse tedavi gruplarında

hastalara önemli faydalar sağlamaktadır. Ayrıca klinik çalışmanın her iki kolunun etkinliğinin “plasebo etkisi”nin olası bir kaybı ile azaltılacağı ve bu kaybın “bu amaçla kontrol edildiği” görüşü öne sürülebilir. Ancak olası faydalar bir klinik çalışma kapsamında her bir grup içerisinde azaltıldığında, ilaç ile plasebo arasında potansiyel istatistiksel anlamlı bir fark artık anlamlı olmayabilmektedir. Demek ki, hastalara optimal tedavinin verildiğinin söylenmesinden kaynaklanan herhangi bir terapötik faydanın klinik çalışma koşullarında kaybedilmiş olacağı açıktır.

Bu düşünceler, bir tedaviyi bir başka tedavi ya da plasebo ile, tedavi seçenekleri ile ilişkili kaçınılmaz önyargılardan bağımsız olarak karşılaştırmak için optimal yöntem olmaya devam eden klinik çalışmalarda (124;132), klinik olarak daha faydalı bilgi sağlamak açısından yöntem değişikliklerinin yararlı olabileceğini düşündürmüştür (130;158;159). Yine de, önerilen bu yöntem değişikliklerine rağmen klinik çalışmalar yukarıda sunulan kısıtlamalardan etkilenmeye devam edecektir. Söz konusu kısıtlamaların daha fazla farkında olmak, klinik çalışmaların sonuçlarını ve çıkarımlarını yorumlamakta sağlık profesyonelleri ve toplum açısından faydalı olacaktır.

C. Klinikte herhangi bir romatizmal sorunu olan her hastanın kontrol vizitinde MDHAQ ve RAPID3 ile kantitatif olarak izlenmesi

Şimdi bir romatoloji kliniği ortamında herhangi bir romatizmal sorunu olan her hastanın kontrol vizitinde MDHAQ ve RAPID3 uygulanması suretiyle kantitatif olarak değerlendirilmesi ve izlenmesi için bir stratejiyi inceleyeceğiz. Bu uygulama romatologlara durum ve sonuçtaki değişiklikleri belgeleme imkanı sunmakta ve romatoloji bakımının kalitesini iyileştirmektedir. Bu bölüm önceki yayınları özetlemekte, kaynakları gözden geçirmekte ve özellikle de (100. referans) yorumlamaktadır (160-164).

Öncelikle araştırma için değil, klinik hasta bakımı için tasarlanmış bir anket kullanmak gerekir. Araştırma için tasarlanmış hasta anketleri uzun ve sıkıcı olabilir ve ilgili bilgiyi elde etmek amacıyla klinisyenin “değerlendirme”

yapabilmesi için tasarlanmamıştır. Gerçekten de araştırma anketleri analiz için bir veri merkezine gönderilmektedir ve hasta bakımına bir fayda katmadığı gibi, kliniğe de ilave bir yük getirebilmektedir (96-98). Buna karşın klinik için tasarlanmış basit hasta anketleri kısadır, klinisyene zaman kazandırabilmekte ve hasta vizitlerinin kalitesini arttırabilmektedir (96-98).

HAQ, yani RA için en yaygın olarak bulunan anket bile, hastalar tarafından kolaylıkla 5-10 dakika içinde doldurulabilirken, pek çok romatolog tarafından günlük rutinde değerlendirilmesi zordur (bazı dikkate değer istisnalar hariç) (108). HAQ' ta skorlama için 40 saniye gerekmektedir. Buna karşın MDHAQ (81;82) yukarıda da belirtildiği üzere RAPID3'ün rutin bakımda 10 saniyeden daha az bir sürede skorlanmasını kolaylaştırmaktadır (165).

MDHAQ'ın (Şekil 1-A) bir sayfası günlük yaşamın 10 aktivitesini, psikolojik stresi değerlendirmeye yönelik 3 madde, ağrı ve global durum için görsel analog skalalar (VAS) ve bir Romatoid Artrit Hastalık Aktivite İndeksi (RADAI) hastanın kendisinin bildirdiği eklem sayımını (16) (RAPID3 skoruna dahil değil) içermektedir. 21 halkalı bir VAS formatı hızlı bir "değerlendirmeyi" ve fiziksel fonksiyon ile RAPID3 için skorlama şablonları skorlamayı kolaylaştırmaktadır (166). Arka sayfası (Şekil 1-B) sistemlerin gözden geçirilmesi, yorgunluk VAS, yakın geçmişteki tıbbi olaylar ve skorlamaya dahil olmayan ancak klinik bakımdan faydalı veriler sağlayan demografik verileri içermektedir.

Sağlık personelinin hasta anketlerinin önemi konusunda bilgilendirmek gerekmektedir. Bu prosedür karmaşık görünebilir, fakat hastadan hekime bilgi akışını kantitatif veriler aracılığı ile düzenleyerek zamandan tasarruf sağlayabilmektedir. Ofis çalışanı romatolog da klinik sonuçlar bakımından ankete önem verdiğini gördüğünde, o da ankete önem vermeye başlar. Ancak anketler hastaya umursamaz bir tavır ile verilir ve/veya araştırma, belgeleme, masrafların iadesi, meslektaşlar ile işbirliği amacı ya da münferit hasta için daha iyi bir bakım dışındaki herhangi başka bir neden ile açıklanır ise personel (ve hastalar) ilgilerini yitirmekte ve gereksiz görecekları bu işten soğumaktadırlar.

Pek çok romatolog hasta anketlerinin yalnızca RA'li hastalar gibi belirli hastalıklar için ya da her 6 ayda bir gibi belirli aralıklarla kullanılabileceğini önermiş olsa da bu yaklaşım, personelin anketleri selektif olarak dağıtılmasını organize etmesinin fiilen mümkün olmaması nedeniyle genellikle günlük hasta bakımında başarısız olmaktadır. Yukarıda da belirtildiği gibi bir MDHAQ her türlü romatizmal hastalığı olan hasta açısından aydınlatıcıdır. Hasta viziti için bir neden var ise, anket için de bir neden vardır.

Anket vizitten sonra değil, romatoloğu görmeden önce, muayene odasından ziyade bekleme odasında doldurulmalıdır. Çoğu hasta romatoloğu görmeden önce bekleme odasında en az 10 dakika ve hatta çoğu zaman daha uzun bir süre geçirmektedir. Bu zaman dilimi bir hasta açısından bir anket doldurmak için en uygun ve en arzu edilen zaman dilimidir. Anket elbette muayene odasında ve hatta muayene sonrasında da doldurulabilir. Muayeneye girmeden önce doldurulması hasta sorunlarına odaklanmaya yardımcı olmakta ve hekime bakım esnasında klinik kararları yönlendirmeye yardımcı olmak açısından bilgi sağlamaktadır. Etkin bir biçimde işleyen bir muayenehane, hastalara 10 dakika erken randevu vererek hasta anketinin doldurulmasına zaman tanıyabilir.

Sağlık profesyonellerinin çoğu kendileri tarafından toplanan verilerin hastanın kendisinin bildirdiği verilere göre daha doğru ve daha aydınlatıcı olduğuna inansa da, fiziksel fonksiyon, ağrı, yorgunluk ve global durumu ilgilendiren pek çok veri sağlık profesyonellerinden ziyade hasta anketi ile daha doğru bir şekilde saptanabilmektedir (79). Hasta anketi doldurduğunda ise yalnızca tek bir gözlemci söz konusudur. Verilerin elde edilmesine bir sağlık profesyoneli dahil ediliyor ise, iki gözlemci söz konusudur. İkinci bir gözlemcinin dahil edilmesi verilerin tekrarlanabilirliğini arttırmak yerine azaltmaktadır. Ayrıca pek çok hasta fonksiyonel maluliyet, ağrı, yorgunluk vs. ile ilgili sorunları olduğundan daha az bildirerek hekimlerini memnun etmeye çalışmaktadır. Hastaların yaklaşık % 20'si bir anketi doldurmak için personelin ya da aileden birinin yardımına ihtiyaç duymaktadır ki bu yardım memnuniyetle sağlanmaktadır (167;168). Yine de anket ne kadar hasta tarafından doldurulur ise, daha

doğru ve tekrarlanabilir olma olasılığı da o kadar yüksektir ve personelin zamanı minimum düzeyde harcanmaktadır.

MDHAQ'daki skorlama şablonları klinikte kullanımına katkıda bulunmaktadır. Tüm yeni MDHAQ sürümleri fiziksel fonksiyon, ağrı ve global tahmin (skor 0-10) gibi üç temel veri seti yönteminden RAPID3 skorlarının, bir cetvel, hesaplayıcı, bilgisayar ya da web sitesi olmaksızın skorlanmasını kolaylaştırmak amacıyla skorlama şablonları içermektedir. Günlük yaşamın 10 aktivitesi 0-10 skorlama şablonu kullanılarak hızlı bir şekilde hesaplanabilmektedir. VAS, cetvelsiz skorlamayı kolaylaştırmak amacıyla geleneksel 10 cm skalasından ziyade 21 adet halka içermektedir (166). Toplam RAPID3 skoru 0-30 olabilir ya da bir 0-10 skoru için 32'ye bölünebilir.

Yukarıda belirtildiği gibi MDHAQ'ın bir hekim tarafından genellikle hasta ile birlikte "değerlendirmesi" hasta ziyaretlerinin kalitesini ve etkinliğini iyileştirmektedir. Böyle bir gözden geçirme için ayrılacak 5 saniye çoğu zaman 2-5 dakikalık bir sorgulamayı gerektirecek bilgiyi sağlamaktadır.

Klinisyene skorları bir sayfa üzerinde düzenli olarak değiştirme imkanı sağlayan bir formatta düzenlenen, hasta anket verileri, seçilmiş laboratuvar testleri ve ilaçları kapsayan iş akış şemaları (Şekil 2) çok faydalıdır. Bu bilgi hastanın seyrine genel bir bakış sağlamaktadır ve maliyet etkin olan bir prosedürdür. Hastaları tek tek izlemek amacıyla iş akış şemaları geliştirmek için bir bilgisayar, bir araç ya da sistem kullanmaya gerek yoktur. Bilgisayar desteği, gruplar halinde hastaların analizleri ve raporları için gereklidir. Verilerin otomasyonu romatolog ve personelin işlerini kolaylaştırmak için gereklidir.

Klinik araştırma içinde tüm hastaların % 80'inde tam veri elde etmek, % 5 hastanın % 100'de veri elde etmeye yeğlenmelidir (98). Bu nedenle alışılmış klinik bakımda uygun ve uygulanabilir olan daha az kapsamlı bir yöntem hiç kantitatif yöntem kullanılmamasına tercih edilmelidir. Bunun yanında bir RAPID3 skoru verilerin "% 80'inden fazlasını" sağlayabilir ve gerçekten de hasta ve hekimin tedavi hedeflerini

şış ve hassas eklem sayıları kadar doğru yansıtarak, hasta değerlendirmesi ve izlenmesi açısından DAS veya CDAI kadar aydınlatıcı olabilir.

Hasta tarafından doldurulan hasta anket skorlarına dayanan RAPID3 skorları bir hasta ziyaretinden diğerine hasta durumu açısından aydınlatıcı kantitatif veri sağlamaktadır. Kantitatif veriler kayıt altına alınacak olur ise, hasta bakımında artmış etkinliğin yanı sıra belgeleme ve daha rasyonel bir izlem elde edilmiş olmaktadır. Herhangi bir veri kaydedilmez ise, bu şans yitirilmekte ve asla yerine konamamaktadır. Tüm romatologların günlük bakımda tüm ziyaretlerde hastalarından bir MDHAQ doldurmalarını istemelerinin ve bir RAPID3 skorlaması yapmalarının (kendileri ya da bir personel tarafından) önemli olduğunun altını çizirim.

Tablo 1. Romatoid artrit (RA) kantitatif yöntemlerin avantajları ve dezavantajları

EKLEM SAYIMLARI	
Avantajları	Dezavantajları
1. RA'ın değerlendirilmesinde en özgün yöntem 2. Klinik çalışmalarda en önemli yöntem 3. Patogenetik mekanizmaları yansıtır	1. Resmi çalışmaların tekrarlanabilirliği zayıftır 2. Gerçekleştirilmesi sıkıcı - hasta vizitinin bölür 3. Eklem sayımlarının hasta anket skorlarına kıyasla plasebo ile iyileşmesi daha olasıdır - zayıf etkinlik 4. Şiş ve hassas eklem sayımı 5-15 yıl içerisinde iyileşirken, radyografik ve fonksiyonel skorlar progresif hasara işaret eder 5. Çoğu romatolog viziti resmi bir eklem sayımını kapsamaz (fakat genellikle dikkatli bir eklem muayenesini kapsar)
RADYOGRAFİLER	
Avantajları	Dezavantajları
1. Mükemmel kantitatif skorlama sistemleri - Sharp, van der Heijde, Larsen, Genant 2. Erozyonlar RA'de patognomonik bulgular gibidir 3. Kümülatif hasta hasarını yansıtır 4. Romatologa iş yükü getirmez	1. Kantitatif skorlamanın gerçekleştirilmesi sıkıcı 2. Tedavi erozyondan önce başlatılmalı - MRI, ultrason daha duyarlı 3. Radyografik hasarın çalışma maluliyeti, ölüm ve hatta eklem replasmanı açısından çok az değeri var
LABORATUVAR TESTLERİ:	
Avantajları	Dezavantajları
1. Çoğu klinik durumda, örneğin kolesterol, hemoglobin, kreatinin, glikoz vs. en önemli yöntem 2. Pek çok test değeri olabilir - CBC, ESR, CRP, RF, anti-CCP 3. Romatologa iş yükü getirmez	1. ESR ve CRP - %40'ında takdim sırasında normal 2. Anti-CCP ve romatoid faktör - hastaların %20-50'sinde negatif 3. Klinik iyileşmeye rağmen, pek çok hastada değişmemiş olabilir 4. Tedavi kararları öncelikle klinik kriterlere dayanır 5. Laboratuvar testleri radyografik hasar açısından yüksek değere sahiptir, ama çalışma maluliyeti veya ölüm açısından zayıf prognostik değere sahiptir.
HASTA ANKETLERİ:	
Avantajları	Dezavantajları
1. Eklem sayımları, ESR, radyografik skorlar, fiziksel yöntemler ile anlamlı korelasyon 2. Geleneksel eklem sayımları, ESR, radyografik skorlar, fiziksel yöntemlere kıyasla plasebo ile daha tekrarlanabilir ve iyileşmesi daha az olası 3. Klinik çalışmalarda ACR 20, 50, 70 ya da DAS kadar aydınlatıcı 4. Çalışma maluliyeti, masraflar, eklem replasmanı ve erken ölümü geleneksel eklem sayımları, radyografiler ve laboratuvar testlerine göre daha iyi öngörür 5. Hasta zaman içerisinde kendi "kontrol"ü olarak görev alabilir	1. RA ile ilgili olmayan non-spesifik gelişmelere bağlı olarak iyileşebilir - progresif hasar ile iyileşir 2. Kültürel ve dil farklılıklarına tabi 3. Bazı hastalar tarafından ciddiye alınmayabilir
GLOBAL YÖNTEMLER:	
Avantajları	Dezavantajları
1. Basit, süratle tayin edilir 2. 0-10 skala değişikliğe duyarlı 3. Klinik çalışmalarda aktif tedaviyi kontrol tedavisinden ayırma etkinliği eklem sayımları, laboratuvar testleri, anket skorlarına eş ya da daha fazla	1. Mutlaka hastalık durumundaki iyileşmelere özgün olmaz 2. Fonksiyonel sınıf gibi American Rheumatism Association (ARA) [artık American College of Rheumatology (ACR)] 1-4 skalası, zaman içerisinde değişikliğe duyarlıdır
ENDEKSLER:	
Avantajları	Dezavantajları
1. Münferit hastaları izlemek için en iyi yöntem 2. Klinik ortamlarda hasta yöntemlerinin RAPID3 indeksleri kolaylıkla hesaplanır	1. Uzun dönem prognozu açısından fonksiyonel durum ile ilgili hasta anketleri kadar değerli olmayabilir 2. DAS, CDAI hassas ve şiş eklem sayımları gerektirir - alışılmış klinik uygulamada genellikle yoktur

Tablo 2: Randomize kontrollü klinik çalışmaların çağdaş klinik araştırmalardaki pragmatik sınırlamaları

1) Dışlama kriterleri bir klinik çalışma için uygun olduğu düşünülen belirli bir tanı almış hastaların uygunluğunu %10'dan daha azı ile sınırlandırabilir
2) Kronik hastalıklarda klinik çalışmaların oldukça kısa olan süresi
3) İstatistiksel güç açısından yetersiz hasta sayısı nedeniyle, klinik olarak önemli farklılıklar istatistiksel olarak anlamlı olmayabilir
4) İstatistiksel anlam mutlaka klinik anlam ya da önemi göstermemektedir
5) Hastanın bir ilaca karşı başka bir ilaca mı yoksa plaseboya mı randomize edildiğinin dışında sonuçları etkileyecek önemli değişkenler klinik çalışma yazılarında genellikle gözardı edilmektedir
6) Esnek olmayan dozaj ve çalışmadan çıkarma kriterleri
7) Hastalık sonuçları açısından optimalin altında belirteçler olabilen belirteç gereksinimi
8) Çoğu çalışmada nadir advers olaylar tanımlanamamaktadır

Şekil 1. Çok boyutlu sağlık değerlendirme anketi (MDHAQ). MDHAQ'nun (A) ön sayfası fonksiyon için 10 aktivite ile ağrı ve global durum tahmini için iki görsel analog skalası (VAS) ve bir romatoid artrit hastalık aktivite indeksi (RADAI) "kendin bildir" eklem sayısı içermektedir. Bu yöntemler için bir skorlama şablonu sayfanın sağ alt kısmında yer almaktadır. Üç hasta tarafından bildirilen yöntemin bir indeksi, hasta indeksi verilerinin rutin değerlendirmesi (RAPID3) MDHAQ'tan 10 saniyeden daha az bir sürede hesaplanabilmektedir. Arka sayfa (B) sistemlerin gözden geçirmesi, yorgunluk VAS, son tıbbi olaylar ve demografik verileri (skorlamaya dahil değildir ancak klinikte faydalı veriler sağlamaktadır) içermektedir.

Şekil 2. Rutin romatoloji kliniğinde hastanın dikey değerlendirilmesini kolaylaştıracak iş akış şeması. İş akış şeması 61 yaşında romatoid artritli bir hastayı göstermektedir. 4 Kasım 2003'te fiziksel fonksiyon skoru 3.3, ağrı skoru 9.5, global durum skoru 9.5 ve RAPID3 skoru 22.3 olarak başvurmuştur (0-30 skalası üzerinde). 10 mg/hafta metotreksat ve 3 mg/gün prednizon ile tedavi edilmiştir. İki ay sonra, 13 Ocak 2004'te RAPID3 skoru 1 olmuştur, bu da remisyona yakın bir duruma işaret etmektedir. 20 Temmuz ve 28 Eylül 2004 ziyaretlerinde belgelendiği üzere neredeyse bir yıl süre ile oldukça iyi olmuştur (RAPID3 skoru 20 Temmuz'da 5.5'tir fakat bu inflamasyona değil, akut sırt zedelenmesine bağlı ortaya çıkmıştır ve tedavisi değiştirilmemiştir).

28 Aralık 2004'te ağır bir alevlenme ile gelmiştir. Eklemleri yine şişmiş ve fiziksel fonksiyon skoru sıfır olmasına rağmen ağrısı 6.0 ve global ise 5.5 olarak bulunmuştur. Hastaya bir anti-TNF ajan, adalimumab, önerilmiş ve hasta almayı kabul etmiştir. İki ay sonra 5 Şubat 2005'te tüm skorları 0'dır, bu da mükemmel bir yanıtı işaret etmektedir. 28 Mart 2006'daki vizitinde gösterildiği üzere bu durum bir yıldan uzun bir süre muhafaza edilmiştir.

S & C

Şekil 1 A

A) Çok boyutlu sağlık değerlendirme anketi (R785-NP2)

Bu anket kan testlerinden, röntgen filmlerinden ya da sizin dışınızdaki herhangi başka bir kaynaktan elde edilemeyen bilgileri içermektedir. Lütfen sizinle şu anda ilgisi olmadığını düşünseniz bile tüm soruları yanıtlamaya çalışın. Yapabildiğiniz kadarını kendiniz doldurmaya çalışın, yardıma ihtiyacınız olduğunda da yardım isteyebilirsiniz. Doğru ya da yanlış cevaplar yoktur. Lütfen tam olarak düşündüğünüz ya da hissettiğiniz gibi cevaplayın. Teşekkürler.

1. Lütfen şu anda kabiliyetiniz ile ilgili en doğru BİR cevabı (✓) ile işaretleyiniz:

SON BİR HAFTA İÇERİSİNDE, aşağıdakileri yapabildiniz mi:	HERHANGİ BİR zorluk olmaksızın	BİRAZ zorlukla	BÜYÜK zorlukla	YAPAMADIM
a. Giyinebildiniz mi, ayakkabılarınızı bağlamak ve düğme iliklemek dahil?	_____0	_____1	_____2	_____3
b. Yatağa girebildiniz mi ve yataktan çıkabildiniz mi?	_____0	_____1	_____2	_____3
c. Dolu bir fincan ya da bardağı ağzınıza götürebildiniz mi?	_____0	_____1	_____2	_____3
d. Dışarıda düz zemin üzerinde dolaşabildiniz mi?	_____0	_____1	_____2	_____3
e. Tüm vücudunuzu yıkayıp kurulayabildiniz mi?	_____0	_____1	_____2	_____3
f. Yerden giysilerinizi kaldırmak için eğilebildiniz mi?	_____0	_____1	_____2	_____3
g. Normal muslukları açıp kapatabildiniz mi?	_____0	_____1	_____2	_____3
h. Bir araba, otobüs, tren ya da uçağa binip inebildiniz mi?	_____0	_____1	_____2	_____3
i. İsteddiğinizde iki mil ya da üç kilometre yürüyebildiniz mi?	_____0	_____1	_____2	_____3
j. İsteddiğinizde, eğlenceli aktivitelere ve spora arzu ettiğiniz kadar katılabildiniz mi?	_____0	_____1	_____2	_____3
k. İyi bir gece uykusu uyuyabildiniz mi?	_____0	_____1.1	_____2.2	_____3.3
l. Endişe duyguları ile ya da sinirli olmakla başa çıkabildiniz mi?	_____0	_____1.1	_____2.2	_____3.3
m. Depresyon duyguları ya da hüznülmakla başa çıkabildiniz mi?	_____0	_____1.1	_____2.2	_____3.3

1.ai FN (0-10):

☐	1=0.3	5=1.7	9=3.0	13=4.3	17=5.7	21=7.0	25=8.3
☐	2=0.7	6=2.0	10=3.3	14=4.7	18=6.0	22=7.3	26=8.7
☐	3=1.0	7=2.3	11=3.7	15=5.0	19=6.3	23=7.7	27=9.0
☐	4=1.3	8=2.7	12=4.0	16=5.3	20=6.7	24=8.0	28=9.3
☐							30=10

2. Durumunuz nedeni ile SON BİR HAFTA İÇERİSİNDE ne kadar ağrınız oldu?

Lütfen aşağıda ağrınızın ne kadar şiddetli olduğunu belirtin:

AĞRI ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ OLABİLDİĞİNCE
YOK 0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10 KÖTÜ BİR AĞRI

PN

☐

Şekil 1 A (devam)

3. Aşağıda listelenen eklem bölgelerinde bugün ne kadar ağrınız olduğunu göstermek için lütfen uygun olan yerleri (✓) ile işaretleyin:

	Yok	Hafif	Orta	Şiddetli		Yok	Hafif	Orta	Şiddetli
a. SOL PARMAKLAR	☐ 0	☐ 1	☐ 2	☐ 3	i. SAĞ PARMAKLAR	☐ 0	☐ 1	☐ 2	☐ 3
b. SOL BİLEK	☐ 0	☐ 1	☐ 2	☐ 3	j. SAĞ BİLEK	☐ 0	☐ 1	☐ 2	☐ 3
c. SOL DİRSEK	☐ 0	☐ 1	☐ 2	☐ 3	k. SAĞ DİRSEK	☐ 0	☐ 1	☐ 2	☐ 3
d. SOL OMUZ	☐ 0	☐ 1	☐ 2	☐ 3	l. SAĞ OMUZ	☐ 0	☐ 1	☐ 2	☐ 3
e. SOL KALÇA	☐ 0	☐ 1	☐ 2	☐ 3	m. SAĞ KALÇA	☐ 0	☐ 1	☐ 2	☐ 3
f. SOL DİZ	☐ 0	☐ 1	☐ 2	☐ 3	n. SAĞ DİZ	☐ 0	☐ 1	☐ 2	☐ 3
g. SOL AYAK BİLEĞİ	☐ 0	☐ 1	☐ 2	☐ 3	o. SAĞ AYAK BİLEĞİ	☐ 0	☐ 1	☐ 2	☐ 3
h. SOL AYAK PARMAKLARI	☐ 0	☐ 1	☐ 2	☐ 3	p. SAĞ AYAK PARMAKLARI	☐ 0	☐ 1	☐ 2	☐ 3
q. BOYUN	☐ 0	☐ 1	☐ 2	☐ 3	r. SIRT	☐ 0	☐ 1	☐ 2	☐ 3

RAPID 3 (0-30):

3. a-PRADAI PTJTCT (0-10):

1=0.2 6=1.3 11=2.3 16=3.3 21=4.4 26=5.4 31=6.4 36=7.5 41=8.5 46=9.6
 2=0.4 7=1.5 12=2.5 17=3.5 22=4.6 27=5.6 32=6.7 37=7.7 42=8.8 47=9.8
 3=0.6 8=1.7 13=2.7 18=3.8 23=4.8 28=5.8 33=6.9 38=7.9 43=9.0 48=1.0
 4=0.8 9=1.9 14=2.9 19=4.0 24=5.0 29=6.0 34=7.1 39=8.1 44=9.2
 5=1.0 10=2.1 15=3.1 20=4.2 25=5.2 30=6.3 35=7.3 40=8.3 45=9.4

MDGLOBAL

(0-10):

4. Hastalığın ve sağlık koşullarının sizi şu anda etkileyebileceği tüm durumları düşünerek, lütfen aşağıda nasıl olduğunuzu belirtin:

ÇOK 0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10 KÖTÜ

Lütfen devam edin

Yalnız Doktor Kullanımı İçin: Pas ya da RAPID 3 skoru 0-30'dan 0-10'a dönüştürme

NR: 1=0.3, 2=0.7, 3=1.0 LS: 4=1.3, 5=1.7, 6=2.0

MS: 7=2.3, 8=2.7, 9=3.0, 10=3.3, 11=3.7 12=4.0

HS: 13=4.3, 14=4.7, 15=5.0, 16=5.3, 17=5.7, 18=6.0, 19=6.3, 20=6.7, 21=7.0, 22=7.3, 23=7.7, 24=8.0, 25=8.3, 26=8.7, 27=9.0, 28=9.3, 29=9.7, 30=10.0

PTGL

Şekil 1 B

5. Son bir ay içerisinde aşağıdakilerden herhangi birini yaşadysanız, lütfen (o) ile işaretleyin:

☐ Ateş	☐ Boğazınızda şişlik	☐ Kollarda veya bacaklarda paralizi
☐ Kilo artışı (>5 kg)	☐ Öksürük	☐ Kollarda veya bacaklarda uyuşma veya titreme
☐ Kilo kaybı (<5 kg)	☐ Nefes darlığı	☐ Bayılma nöbetleri
☐ Hasta hissetmek	☐ Hırıltı	☐ Ellerde şişme
☐ Baş ağrıları	☐ Göğüste ağrı	☐ Ayak bileklerinde şişme
☐ Alışılmadık yorgunluk	☐ Kalp vurumu (palpitasyonlar)	☐ Diğer eklemlerde şişme
☐ Şişmiş bezeler	☐ Yutkunma güçlüğü	☐ Eklem ağrısı
☐ İştah kaybı	☐ Kalp yanması ya da mide gazı	☐ Sırt ağrısı
☐ Ciltte döküntü veya ürtiker	☐ Mide ağrısı ya da kramplar	☐ Boyun ağrısı
☐ Alışılmadık morarma ya da kanama	☐ Bulantı	☐ Dükkânlarda satılmayan ilaç kullanımı
☐ Diğer cilt sorunları	☐ Kusma	☐ Sigara içmek
☐ Saç dökülmesi	☐ Kabızlık	☐ Günde ikiden fazla alkollü içecek
☐ Kuru göz	☐ İshal	☐ Depresyon-hüzünlü hissetme
☐ Diğer göz sorunları	☐ Koyu renkli ya da kanlı gaita	☐ Endişe-sinirli hissetme
☐ İştme ile ilgili sorunlar	☐ İdrara çıkma ile ilgili sorunlar	☐ Düşünme ile ilgili sorunlar
☐ Kulaklarda çınlama	☐ Jinekolojik sorunlar	☐ Hafıza ile ilgili sorunlar
☐ Tıkalı burun	☐ Baş dönmesi	☐ Uyku ile ilgili sorunlar
☐ Ağızda yaralar	☐ Denge kaybı	☐ Cinsel sorunlar
☐ Ağız kuruluğu	☐ Kas ağrısı, acısı, ya da krampları	☐ Cinsel organlarda yanma
☐ Koku ya da tat alma duyusu ile ilgili sorunlar	☐ Kas güçsüzlüğü	☐ Sosyal aktivitelerle ilgili sorunlar

5. ROS

☐
6. SON BİR HAFTA İÇERİSİNDE sabah uyanığınızda tutulmuş hissettiniz mi? ☐Evet ☐Hayır

Eğer "hayır" cevabı verdiyseniz, lütfen 7. maddeye geçin. Eğer "evet" cevabı verdiyseniz gün içinde olabileceğiniz kadar hareketli olmanıza kadar geçen süre için dakikaların..... ya da saatlerin..... sayısını belirtin.

6. AM

☐

Şekil 1 B (devam)

10. Son 6 ayda aşağıdakilerde hangisini yaşadınız: (Lütfen (✓) ile işaretleyin):

- | | |
|--|---|
| ☐ Hayır ☐ Evet Ameliyat | ☐ Hayır ☐ Evet Artrit ilaçlarında ya da diğer ilaçlarda değişiklik(ler) |
| ☐ Hayır ☐ Evet Hastane yatışı | ☐ Hayır ☐ Evet Adres değişikliği (değişiklikleri) |
| ☐ Hayır ☐ Evet Yeni bir hastalık, kaza ya da travma | ☐ Hayır ☐ Evet Medeni durumda değişiklik(ler) |
| ☐ Hayır ☐ Evet Önemli yeni bir semptom | ☐ Hayır ☐ Evet İş ya da görev değişikliği, işten ayrılma, emeklilik |
| ☐ Hayır ☐ Evet Herhangi bir ilacın yan etkisi (etkileri) | ☐ Hayır ☐ Evet Sağlık sigortasında değişiklik |
| ☐ Hayır ☐ Evet Düzenli sigara içmek | ☐ Hayır ☐ Evet Primer bakımda değişiklik ya da başka doktor |

Lütfen tüm "evet" cevaplarını aşağıda açıklayın ya da sizi etkileyen diğer tüm sağlık konularını belirtin:

CİNSİYET: Kadın, Erkek ETNİK GRUP: Asyalı, Siyah, Hispanik, Beyaz, Diğer _____

Mesleğiniz: _____ Lütfen tamamladığınız eğitim süresini (yıl) daire içine alın:

İş Durumu: ☐ Tam zamanlı ☐ Yarı zamanlı ☐ Engelli 1 2 3 4 5 6 7 8 9 10☐ Evde çalışıyor ☐ Kendi işi ☐ Emekli 11 12 13 14 15 16 17 18 19 20☐ İş arıyor ☐ Diğer _____ Lütfen kilonuzu yazın: _____ lbs. Boy: _____ inç

Adınız _____ DoğumTarihiniz _____ Bugünün Tarihi _____

Tıbbi bakımınızı takip edebilmemize yardımcı olmak için bu anketi doldurduğunuz için teşekkürler.

YALNIZCA DOKTOR KULLANIMI İÇİN: Anketi aldım (✓) ile işaretleyin): _____ Tarih _____

İsim: _____ İmza _____

PT İsmi _____ DX ICD9 _____ 710.9 _____, Başlangıç (ay/yıl) _____ 09/2003 _____ Kayıt# _____

Romatolog _____ TP _____, 1. ziyaret (ay/yıl) _____ 4 Ksm 03, RF: Poz / Neg + ise, titrasyonu _____, ANA: Poz / Neg + ise, titrasyonu _____

Adres _____, şehir _____, SOKAK, POSTA KODU Ev tel _____

SSN# _____, DOB _____ /- /1942, Cinsiyet M / F, Medeni durum: _____ M _____, İrk: _____ C _____

İş durumu: _____ FT _____, Meslek _____ #Eğitim Süresi 12, Verilen Onam: Y/ N, 1° DR _____ DR Tel _____

R785NP2

10. XM



M
D
H
A
QL
A
B
O
R
A
T
U
Rİ
L
A
Ç
L
A
R

Şekil 2	4 Ksm03	13Ocak04	20Tem04	28Eyl04	28Arl04	08Şbr05	28Mar06
TARİH							
FONKSİYONEL DURUM (FN) (0-10)	3.3	0	0	0	0	0	0
AĞRI (PN) (0-10)	9.5	0.5	3.5	0.5	6.0	0	0.5
HASTAGLOBAL (PTGL) (0-10)	9.5	0.5	2.0	1.0	5.5	0	0.5
RAPID 3 (0-30)	22.3	1.0	5.5	1.5	11.5/4.0	0/0	1.0/0.3
PT EKLEM SAYIMI (JT CT) (0-10)					1.9	0	0
RAPID 4 (0-40)					13.4/3.5	0/0	1.0/0.3
HEKİM GLOBAL (MDACT) (0-10)					6.5	1.0	0.5
RAPID 5 (0-50)					19.9/4.0	1.0/0.02	1.5/0.4
KİLO (lbs)	167	167	163.8	159	168	166	117
KAN BASINCI (mm/Hg)	114/70	131/81	116/76	128/80	111/71	120/72	129/79
ESR (mm/hr) (M:0-20/F:0-30)	43	8	11	10	14	14	14
CRP (mg/dL) (0-10)	30	3		7	6	8	9.3
WBC (Thou/uL) (4-11)	6.3	7.9	7.1	8.1	9.1	9.6	9.4
HGB (g/dL-M:14/F:12) VEYA hct (%) (M:42/F:37)	16.8	17	15.9	16.1	16.6	17	15.3
TROMBOSİTLER (thou/uL) (150-400)	179	207	184	203	207	177	193
ALBUMİN (g/dL) (3.5-5.0)	3.9	4.1	4.4	4	4.4	4.6	4.1
SGOT (U/L) (4-40) VEYA SGPT (U/L) (4-40)	18	17	22	18	20	32	21
KREATİNİN (mg/dL) (0.7-1.5)	1.1	0.8	0.9	0.9	0.9	1.1	1.0
MED KODLARI: N-yeni ilaç, O-vizitte açık, X-toksosite, C-doz değişikliği, D-kesme, T-azalma, R-geri dönme, I-enjeksiyon, V-yalnızca bugün							
Naproksen	0-880 Q6H	440 BID	440 BID	440 BID	440 BID	D-440 BID	
Ranitidin	0-150 BID	150 BID	150 BID	150 BID	150 BID	150 BID	75 BID
Kodeinli Asetaminofen	0-30 TID	30 TID	D-30 TID				
Prednizon	N-3 QD	1 BID	C-4 BID	C-3 BID	T-3 BID	T-2 BID	C-5 QD
Metotreksat	N-10 QWK	20 QWK	C-15 QWK	15 QWK	C-25 QWK	15 QWK	15 QWK
Folik Asit	N-1 QD	1QD	1 QD	1 QD	1 QD	1 QD	1 QD
Adalimumab					N-40 QOW	40 QOW	40 QOW
Depo-Medrol					V-80		

S. Ben bir genetikçiyim. Basit bir sorum olacak. Cevabından emin değilim. Bu tür Cox-bazlı analizde kantitatif veri kullanılması her zaman mümkün olmamaktadır; özellikle romatoid artrit gibi hastalıklarda çünkü bu kantitatif verilerin oldukça geniş ve yüksek güvenlik intervalleri vardır ve 1+1 eşittir 2 değildir. Tüm bu etkileşimler dinamikdir. Yani etkileşimler dinamik ise nasıl öngöründe bulunabiliyorsunuz?

C. Sanıyorum yöntemlerimizin çok ham olduğunu ve daha iyi yöntemlere ihtiyacımız olduğunu söylüyorsunuz. Buna kesinlikle katılıyorum. Son 20 yıldır elimizde olan en iyi yöntemleri kullanarak oldukça büyük ilerlemeler kaydettik. Hasta anketinin klinik kullanımı ile bir "biomarker" doğru tanı geliştirme kapasitesine katkıda bulunacaktır. Belki sizin gibi hekimler bunu gelecek 10 yıl içinde bulacaklardır ve belki de biz bunları dinlemek üzere böylesine güzel bir mekâna yeniden davet edileceğiz. Bu kesinlikle bir gelişme olurdu.

Ancak ben, romatoid faktörün yapısını anladığımızda, romatoid artritin ortadan kalkacağına inanıldığı 60'lı ve 70'li yılları hatırlayacak kadar yaşıyım. Bu açıdan bakıldığında, herhangi bir klinik ölçüm yapmamıza gerek yoktu. Bu büyük bir hataydı zira romatoid artritin uzun dönem sonuçlarına dair büyük yanılgılar yarattı. Benim gibi kır saçlı profesörler tarafından remisyon indükleyici tedavi olduğu öğretilmekteydi. Bu durum klinik deneylerde bir yıl süre ile geçerliydi, ancak zaman içerisinde durum değişti.

Sizin söylediğinize katılmıyor değilim fakat tedavilerimizin etkili olup olmadığını bulmak konusunda hastalarımıza karşı da entelektüel ve etik bir sorumluluğumuz var; sadece profesörlerimizin bize söylediklerini takip etme yükümlülüğümüz yok. Bir anlamda bugünün konusu da işte bu. Mevcut olan en iyi yöntemleri kullanmalıyız, fakat bu bizi daha iyi yöntemler aramaktan alıkoymamalı.

S. Bu anketlerin kantitatif veriler olduğunu söylüyorsunuz. Benim sorum anket verilerinin ne kadar kantitatif olduğu. Hastanın eğitiminin kısıtlılığı ile sonuç arasındaki ilişki mi? Genel olarak anketlerde

yanıt olmaması gibi büyük bir sorun. Bununla nasıl başa çıkıyorsunuz?

C. Yanıt olmaması basit bir konu. Benim kliniğimde yanıt yok sıfırdır, yani esasen bizim ardışık bir hasta veri tabanımız vardır. Ardışık bir hasta veri tabanı gerçekten de oldukça aydınlatıcı olabilir; klinik deneyler kadar aydınlatıcı. Kayıtlar için bu geçerli değildir; örneğin anti-TNF tedavisi alan hastaların kayıtlarının bir seleksiyon sorunu vardır. Hasta anketi verilerinin tekrarlanabilirliğini tartışalım. Tüm itinalı analizler hasta "kendin bildir" skorlarının doktorların eklem sayımlarından ya da radyografik skorlardan daha tekrarlanabilir ve daha güvenilir olduğunu göstermektedir. Bunun asıl nedeni, sanıyorum, formu hasta doldurduğunda yalnızca bir gözlemci vardır. Oysa ikinci bir gözlemci, bir doktor ya da bir hemşire ya da bir terapist olsun, dahil olduğunda, iki gözlemci sorunu ortaya çıkmaktadır, zira bu iki gözlemci aynı şeyi aynı şekilde görmezler. Yani bundan kaçınmanızı öneririm. Elbette bir hasta okuma yazma bilmiyor ise, ki bu tüm toplumların bir gerçeğidir, hasta soruları nasıl yanıtlayacağını sorabilir. Ben hastalara daima "doğru" cevabın olmadığını ve "uygun" cevabın kendilerinin verdiği cevap olduğunu söylerim. Yardıma ihtiyaç vardır, tercihen bir sağlık profesyonelinin çok bir aile bireyi tarafından.

Lütfen hasta verilerinin çalışma maluliyeti ya da ölüm sonuçlarını radyografiler ya da laboratuvar testlerinden daha anlamlı bir şekilde öngördüğünü unutmayın. Yani hasta anketlerinin laboratuvar testleri kadar kantitatif olduğunu söyleyebilirim. Biz sadece yalan söylemeyen verileri kaydedebiliriz. Veriler hakkındaki yorumumuz hatalı olabilir. Şu anda hasta anketi verileri elimizdeki en iyi yöntemlerdir.

S. Biz de bu anketlere büyük değer veriyoruz ve zaman zaman hastanın tarif ettiği şekli ile ağrı ve fonksiyonel durum arasında uyumsuzluklar olmaktadır. Sanıyorum siz fonksiyonel durumu baz alıyorsunuz?

C. Ağrılarını abarttığımızı düşünen bazı hastalarımız vardır, ancak ağrısını saklamaya çalışan hastalarımız da vardır. Ağrı skorunu yorumlamak doktorun işidir, tıpkı benim normal ESR'e sahip bu hastayı anti-TNF ile

tedavi ettiğim gibi. Bu bayanın hayatının çok iyi gitmediği benim tespitimdi. Elimin altında kendisini tedavi edecek imkanlar vardı ve ben de bunu yapmaya karar verdim. Eninde sonunda, ölçümün yaklaşımım için bir kilometre taşı olduğunu vurgulamak isterim fakat bir hekim yöntemi yorumlamalı. Klinik yargı karar sürecinin bir parçası olmalı. Belirli bir sıcaklık, ESR ya da herhangi bir klinik yöntem için kurallarımız olamaz. O anda mevcut verilerle hasta için en iyisi neyse, o yapılmalıdır.

REFERANSLAR

1. Goldsmith CH, Smythe HA, Helewa A. Interpretation and power of pooled index. *J Rheumatol* 1993; 20:575-8.
2. van Riel PLCM. Provisional guidelines for measuring disease activity in clinical trials on rheumatoid arthritis (Editorial). *Br J Rheumatol* 1992; 31:793-4.
3. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis Rheum* 1993; 36:729-40.
4. Tugwell P, Boers M. OMERACT Committee. Proceedings of the OMERACT Conferences on outcome measures in rheumatoid arthritis clinical trials, Maastricht, Netherlands. *J Rheumatol* 1993; 20:527-91.
5. van der Heijde DMFM, van't Hof M, van Riel PLCM, van de Putte LBA. Development of a disease activity score based on judgment in clinical practice by rheumatologists. *J Rheumatol* 1993; 20:579-81.
6. Prevoo MLL, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LBA, van Riel PLCM. Modified disease activity scores that include twenty-eight-joint counts: Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995; 38:44-8.
7. Aletaha D, Smolen J. The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005; 23:S100-S108.
8. Cooperating Clinics Committee of the American Rheumatism Association. A seven-day variability study of 499 patients with peripheral rheumatoid arthritis. *Arthritis Rheum* 1965; 8:302-35.
9. American Rheumatism Association. Dictionary of the rheumatic diseases vol. I: signs and symptoms. New York: Contact Associates International; 1982.
10. Decker JL. American Rheumatism Association nomenclature and classification of arthritis and rheumatism. *Arthritis Rheum* 1983; 26:1029-32.
11. Ritchie DM, Boyle JA, McInnes JM, Jasani MK, Dalakas TG, Grieveon P et al. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *Q J Med* 1968; New Series 37:393-406.
12. Egger MJ, Huth DA, Ward JR, Reading JC, Williams HJ. Reduced joint count indices in the evaluation of rheumatoid arthritis. *Arthritis Rheum* 1985; 28:613-9.
13. Fuchs HA, Callahan LF, Kaye JJ, Brooks RH, Nance EP, Pincus T. Radiographic and joint count findings of the hand in rheumatoid arthritis: Related and unrelated findings. *Arthritis Rheum* 1988; 31:44-51.
14. Sokka T, Pincus T. Quantitative joint assessment in rheumatoid arthritis. *Clin Exp Rheumatol* 2005; 23:S58-S62.
15. Mason JH, Anderson JJ, Meenan RF, Haralson KM, Lewis-Stevens D, Kaine JL. The Rapid Assessment of Disease Activity in Rheumatology (RADAR) questionnaire: validity and sensitivity to change of a patient self-report measure of joint count and clinical status. *Arthritis Rheum* 1992; 35:156-62.
16. Stucki G, Liang MH, Stucki S, Brühlmann P, Michel BA. A self-administered rheumatoid arthritis disease activity index (RADAI) for epidemiologic research. *Arthritis Rheum* 1995; 38:795-8.
17. Escalante A. What do self-administered joint counts tell us about patients with rheumatoid arthritis? *Arthritis Care Res* 1998; 11:280-90.
18. Houssien DA, Stucki G, Scott DL. A patient-derived disease activity score can substitute for a physician-derived disease activity score in clinical research. *Rheumatology* 1999; 38:48-52.
19. Taal E, Abdel-Nasser AM, Rasker JJ, Wiegman O. A self-report Thompson articular index: what does it measure? *Clin Rheumatol* 2006;125.
20. Hanly JG, Mosher D, Sutton E, Weerasinghe S, Theriault D. Self-assessment of disease activity by patients with rheumatoid arthritis. *J Rheumatol* 2006;1531.
21. Prevoo MLL, Kuper IH, van Leeuwen M, van't Hof M, van de Putte LBA, van Riel PLCM. Validity and reproducibility of self-administered joint counts. A prospective longitudinal followup study in patients with rheumatoid arthritis. *J Rheumatol* 2006;841.
22. Voon E, Phan L, Ruttkay D, Madera X, McAteer A, Neely K et al. Patient self-administered joint tenderness and swelling counts in RA are reliable and responsive to therapy. *Arthritis Rheum* 2006;S286.
23. Hewlett SE, Haynes J, Shepstone L, Kirwan JR. Rheumatoid arthritis patients cannot accurately report signs of inflammatory activity. *British Journal of Rheumatology* 2006;547.
24. Stewart MW, Palmer DG, Knight RG, Highton J. A self-report articular index: relationship to variations in

- mood and disease activity measures. *British Journal of Rheumatology* 2006;63:1.
25. Wolfe F. Comparative usefulness of c-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. *J Rheumatol* 1997; 24:1477-85.
 26. Pincus T. The DAS is the most specific measure, but a patient questionnaire is the most informative measure to assess rheumatoid arthritis. *J Rheumatol* 2006; 33:834-7.
 27. Bombardier C, Tugwell P, Sinclair A, Dok C, Anderson G, Buchanan WW. Preference for endpoint measures in clinical trials: results of structured workshops. *J Rheumatol* 1982; 9:798-801.
 28. Wolfe F, Pincus T, Thompson AK, Doyle J. The assessment of rheumatoid arthritis and the acceptability of self-report questionnaires in clinical practice. *Arthritis Care Res* 2003; 49:59-63.
 29. Lewis PA, O'Sullivan MM, Rumbold WR, Coles EC, Jessop JD. Significant changes in Ritchie scores. *Br J Rheumatol* 1988; 27:32-6.
 30. Hart LE, Tugwell P, Buchanan WW, Norman GR, Grace EM, Southwell D. Grading of tenderness as a source of interrater error in the Ritchie articular index. *J Rheumatol* 1985; 12:716-7.
 31. Thompson PW, Hart LE, Goldsmith CH, Spector TD, Bell MJ, Ramsden MF. Comparison of four articular indices for use in clinical trials in rheumatoid arthritis: patient, order and observer variation. *J Rheumatol* 1991; 18:661-5.
 32. Klinkhoff AV, Bellamy N, Bombardier C, Carrette S, Chalmers A, Esdaile JM et al. An experiment in reducing interobserver variability of the examination for joint tenderness. *J Rheumatol* 1988; 15:492-4.
 33. Scott DL, Choy EHS, Greeves A, Isenberg D, Kassiror D, Rankin E et al. Standardising joint assessment in rheumatoid arthritis. *Clinical Rheumatology* 2006;57:9.
 34. Wakefield RJ, Kong KO, Conaghan PG, Brown AK, O'Connor PJ, Emery P. The role of ultrasonography and magnetic resonance imaging in early rheumatoid arthritis. *Clin Exp Rheumatol* 2003; 21:S42-S49.
 35. Scott DL, Grindulis KA, Struthers GR, Coulton BL, Popert AJ, Bacon PA. Progression of radiological changes in rheumatoid arthritis. *Ann Rheum Dis* 1984; 43:18-17.
 36. Pincus T, Callahan LF, Sale WG, Brooks AL, Payne LE, Vaughn WK. Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum* 1984; 27:864-72.
 37. Hawley DJ, Wolfe F. Sensitivity to change of the Health Assessment Questionnaire (HAQ) and other clinical and health status measures in rheumatoid arthritis: results of short term clinical trials and observational studies versus long term observational studies. *Arthritis Care Res* 1992; 5:130-6.
 38. Sharp JT, Wolfe F, Mitchell DM, Bloch DA. The progression of erosion and joint space narrowing scores in rheumatoid arthritis during the first twenty-five years of disease. *Arthritis Rheum* 1991; 34:660-8.
 39. Egsmose C, Lund B, Borg G, Pettersson H, Berg E, Brodin U et al. Patients with rheumatoid arthritis benefit from early 2nd line therapy: 5 year followup of a prospective double blind placebo controlled study. *J Rheumatol* 1995; 22:2208-13.
 40. Fex E, Jonsson K, Johnson U, Eberhardt K. Development of radiographic damage during the first 5-6 yr of rheumatoid arthritis. A prospective follow-up study of a Swedish cohort. *Br J Rheumatol* 1996; 35:1106-15.
 41. Mulherin D, Fitzgerald O, Bresnihan B. Clinical improvement and radiological deterioration in rheumatoid arthritis: Evidence that pathogenesis of synovial inflammation and articular erosion may differ. *Br J Rheumatol* 1996; 35:1263-8.
 42. Leirisalo-Repo M, Paimela L, Peltomaa R, Hiekkä K, Möttönen T, Sokka T et al. Functional and radiological outcome in patients with early RA - A longitudinal observational study. *Arthritis and Rheumatism* 42, S130. 1999. Ref Type: Abstract
 43. Graudal N, Tarp U, Jurik AG, Galløe AM, Garred P, Milman N et al. Inflammatory patterns in rheumatoid arthritis estimated by the number of swollen and tender joints, the erythrocyte sedimentation rate, and hemoglobin: Longterm course and association to radiographic progression. *J Rheumatol* 2000; 27:47-57.
 44. Welsing PMJ, van Gestel AM, Swinkels HL, Kiemeneij LALM, van Riel PLCM. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. *Arthritis Rheum* 2001; 44:2009-17.
 45. Puolakkä K, Kautiainen H, Möttönen T, Hannonen P, Korpela M, Hakala M et al. Early suppression of disease activity is essential for maintenance of work capacity in patients with recent-onset rheumatoid arthritis: five-year experience from the FIN-RACO trial.

- Arthritis Rheum 2005; 52:36-41.
46. Puolakka K, Kautiainen H, Mottonen T, Hannonen P, Hakala M, Korpela M et al. Predictors of productivity loss in early rheumatoid arthritis: a 5 year follow up study. *Ann Rheum Dis* 2005; 64:130-3.
 47. Tugwell P, Wells G, Strand V, Maetzel A, Bombardier C, Crawford B et al. Ciovement as reflected in measures of function and health-related quality of life following treatment with leflunomide compared with methotrexate in patients with rheumatoid arthritis: Sensitivity and relative efficiency to detect a treatment effect in a twelve-month, placebo-controlled trial. *Leflunomide rheumatoid arthritis investigators group. Arthritis Rheum* 2000; 43:506-14.
 48. Tugwell P, Wells G, Strand V, Maetzel A, Bombardier C, Crawford B et al. Clinical improvement as reflected in measures of function and health-related quality of life following treatment with leflunomide compared with methotrexate in patients with rheumatoid arthritis: Sensitivity and relative efficiency to detect a treatment effect in a twelve-month, placebo-controlled trial. *Arthritis Rheum* 2000; 43:506-14.
 49. Pincus T, Amara I, Segurado OG, Bergman M, Koch GG. Relative efficiencies of physician/assessor global estimates and patient questionnaire measures are similar to or greater than joint counts to distinguishadalimumab from control treatments in rheumatoid arthritis clinical trials. *J Rheumatol* 2008; 35:201-5.
 50. Pincus T, Segurado OG. Most visits of most patients with rheumatoid arthritis to most rheumatologists do not include a formal quantitative joint count. *Ann Rheum Dis* 2006; 65:820-2.
 51. Steinbrocker O, Traeger CH, Batterman RC. Therapeutic criteria in rheumatoid arthritis. *JAMA* 1949; 140:659-62.
 52. Larsen A. A radiological method for grading the severity of rheumatoid arthritis, Academic dissertation, University of Helsinki. Helsinki: 1974.
 53. Larsen A, Dale K, Eek M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. *Acta Radiol [Diagn] (Stockh)* 1977; 18:481-91.
 54. Sharp JT, Lidsky MD, Collins LC, Moreland J. Methods of scoring the progression of radiologic changes in rheumatoid arthritis: correlation of radiologic, clinical and laboratory abnormalities. *Arthritis Rheum* 1971; 14:706-20.
 55. Sharp JT. Scoring radiographic abnormalities in rheumatoid arthritis. *J Rheumatol* 1989; 16:568-9.
 56. van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 1999; 26:743-5.
 57. Landewe R, van der Heijde D. Radiographic progression in rheumatoid arthritis. *Clin Exp Rheumatol* 2005; 23:S63-S68.
 58. van der Heijde DMFM, van Leeuwen MA, van Riel PLCM, Koster AM, van't Hof MA, van Rijswijk MH et al. Biannual radiographic assessments of hands and feet in a three-year prospective followup of patients with early rheumatoid arthritis. *Arthritis Rheum* 1992; 35:26-34.
 59. Rau R, Wassenberg S, Herborn G, Stucki G, Gebler A. A new method of scoring radiographic change in rheumatoid arthritis. *J Rheumatol* 1998; 25:2094-107.
 60. Rau R, Wassenberg S, Herborn G, Perschel WT, Freitag G. Identification of radiologic healing phenomena in patients with rheumatoid arthritis. *J Rheumatol* 2001; 28:2608-15.
 61. Sharp JT. Radiological Assessment of Joint Damage: The Premier Outcome Measure in Rheumatoid Arthritis. Current Status and Future Potential. In: Wolfe F, Pincus T, editors. *Rheumatoid Arthritis: Pathogenesis, Assessment, Outcome, and Treatment*. New York: Marcel Dekker, Inc.; 1994. 167-89.
 62. Pincus T, Callahan LF. What is the natural history of rheumatoid arthritis? *Rheum Dis Clin North Am* 1993; 19:123-51.
 63. Pincus T, Sokka T. Quantitative measures for assessing rheumatoid arthritis in clinical trials and clinical care. *Best Pract Res Clin Rheumatol* 2003; 17:753-81.
 64. Pincus T, Sokka T. Quantitative measures and indices to assess rheumatoid arthritis in clinical trials and clinical care. *Rheum Dis Clin N Am* 2004; 30:725-51.
 65. Pincus T, Brooks RH, Callahan LF. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. *Ann Intern Med* 1994; 120:26-34.
 66. Engel GL. The need for a new medical model: a challenge for biomedicine. *Science* 1977; 196:129-36.
 67. Rose HM, Ragan C, Pearce E, Lipman MO. Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis. *Proc Soc Exp Biol Med* 1948; 68:1-6.
 68. Schellekens GA, de Jong BAW, van den Hoogen FHJ,

- van de Putte LBA, van Venrooij WJ. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. *J Clin Invest* 1998; 101:273-81.
69. Westergren A. Studies of the suspension stability of the blood in pulmonary tuberculosis. *Acta Med Scand* 1920; 54:247-82.
 70. Kushner I. The phenomenon of the acute phase response. *Ann N Y Acad Sci* 1982; 389:39.
 71. Mallya RK, de Beer FC, Berry H, Hamilton EB, William BE, Pepys MB. Correlation of clinical parameters of disease activity in rheumatoid arthritis with serum concentration of C-reactive protein and erythrocyte sedimentation rate. *J Rheumatol* 1982; 9:224-8.
 72. Riedemann JP, Muñoz S, Kavanaugh A. The use of second generation anti-CCP antibody (anti-CCP2) testing in rheumatoid arthritis – a systematic review. *Clin Exp Rheumatol* 2005; 23:S69-S76.
 73. Sokka T, Pincus T. Most patients receiving routine care for rheumatoid arthritis in 2001 did not meet inclusion criteria for most recent clinical trials or American College of Rheumatology criteria for remission. *J Rheumatol* 2003; 30:1138-46.
 74. Pinals RS, Masi AT, Larsen RA, et al. Preliminary criteria for clinical remission in rheumatoid arthritis. *Arthritis Rheum* 1981; 24:1308-15.
 75. van Gaalen FA, Linn-Rasker SP, van Venrooij WJ, de Jong BA, Breedveld FC, Verweij CL et al. Autoantibodies to cyclic citrullinated peptides predict progression to rheumatoid arthritis in patients with undifferentiated arthritis: a prospective cohort study. *Arthritis Rheum* 2004; 50:709-15.
 76. Wolfe F, Michaud K. The clinical and research significance of the erythrocyte sedimentation rate. *J Rheumatol* 1994; 21:1227-37.
 77. Pincus T, Sokka T. Prevalence of normal erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) on presentation of patients with rheumatoid arthritis (RA) at two rheumatology settings, one in the US and the other in Finland: Is a patient questionnaire a better quantitative measure of clinical severity? *Arthritis and Rheumatism* 52[9], S127. 2005. Ref Type: Abstract
 78. Wolfe F, Pincus T. The level of inflammation in rheumatoid arthritis is determined early and remains stable over the longterm course of the illness. *J Rheumatol* 2001; 28:1817-24.
 79. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23:137-45.
 80. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Med Care* 1992; 30:473-81.
 81. Pincus T, Swearingen C, Wolfe F. Toward a multidimensional health assessment questionnaire (MDHAQ): Assessment of advanced activities of daily living and psychological status in the patient friendly health assessment questionnaire format. *Arthritis Rheum* 1999; 42:2220-30.
 82. Pincus T, Sokka T, Kautiainen H. Further development of a physical function scale on a multidimensional health assessment questionnaire for standard care of patients with rheumatic diseases. *J Rheumatol* 2005; 32:1432-9.
 83. Pincus T, Strand V, Koch G, Amara I, Crawford B, Wolfe F et al. An index of the three core data set patient questionnaire measures distinguishes efficacy of active treatment from placebo as effectively as the American College of Rheumatology 20% response criteria (ACR20) or the disease activity score (DAS) in a rheumatoid arthritis clinical trial. *Arthritis Rheum* 2003; 48:625-30.
 84. Pincus T, Amara I, Koch GG. Continuous indices of Core Data Set measures in rheumatoid arthritis clinical trials: lower responses to placebo than seen with categorical responses with the American College of Rheumatology 20% criteria. *Arthritis Rheum* 2005; 52:1031-6.
 85. Wolfe F, Michaud K, Pincus T. A composite disease activity scale for clinical practice, observational studies and clinical trials: the patient activity scale (PAS/PAS-II). *J Rheumatol* 2005; 32:2410-5.
 86. Hurst NP, Jobanputra P, Hunter M, Lambert M, Lochhead A, Brown H. Validity of Euroqol-a generic health status instrument-in patients with rheumatoid arthritis. *Br J Rheumatol* 1994; 33:655-62.
 87. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975; 1:277-99.
 88. Melzack R. The short-form McGill Pain Questionnaire. *Pain* 1987; 30:191-7.
 89. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989; 46:1121-3.
 90. Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR patient

- preference disability questionnaire - an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *J Rheumatol* 1987; 14:446-51.
91. Clinch J, Tugwell P, Wells G, Shea B. Individualized functional priority approach to the assessment of health related quality of life in rheumatology. *J Rheumatol* 2001; 28:445-51.
 92. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988; 15:1833-40.
 93. Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. *J Rheumatol* 1991; 18:728-33.
 94. Fries JF, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. *Clin Exp Rheumatol* 2005; 23:S53-S57.
 95. Hurst NP, Ruta DA, Kind P. Comparison of the MOS Short Form-12 (SF12) health status questionnaire with the SF36 patients with rheumatoid arthritis. *Br J Rheumatol* 1998; 37:862-9.
 96. Pincus T, Wolfe F. An infrastructure of patient questionnaires at each rheumatology visit: Improving efficiency and documenting care. *J Rheumatol* 2000; 27:2727-30.
 97. Wolfe F, Pincus T. Listening to the patient: A practical guide to self-report questionnaires in clinical care. *Arthritis Rheum* 1999; 42:1797-808.
 98. Pincus T, Wolfe F. Patient questionnaires for clinical research and improved standard patient care: is it better to have 80% of the information in 100% of patients or 100% of the information in 5% of patients? *J Rheumatol* 2005; 32:575-7.
 99. Bergmann FJ, Chassany O, Gandiol J, Deblois P, Kanis JA, Asegregta JM et al. A randomised clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain. *Clin Trials Metaanal* 1994; 29:41-7.
 100. Pincus T, Yazici Y, Bergman M. Saving time and improving care with a multidimensional health assessment questionnaire: 10 practical considerations. *J Rheumatol* 2006; 33:448-54.
 101. Pincus T, Chung C, Segurado OG, Amara I, Koch GG. An index of patient self-reported outcomes (PRO Index) discriminates effectively between active and control treatment in 4 clinical trials of adalimumab in rheumatoid arthritis. *J Rheumatol* 2006; 33:2146-52.
 102. Wolfe F, Cathey MA. The assessment and prediction of functional disability in rheumatoid arthritis. *J Rheumatol* 1991; 18:1298-306.
 103. Callahan LF, Bloch DA, Pincus T. Identification of work disability in rheumatoid arthritis: Physical, radiographic and laboratory variables do not add explanatory power to demographic and functional variables. *J Clin Epidemiol* 1992; 45:127-38.
 104. Wolfe F, Hawley DJ. The longterm outcomes of rheumatoid arthritis: Work disability: A prospective 18 year study of 823 patients. *J Rheumatol* 1998; 25:2108-17.
 105. Sokka T, Kautiainen H, Möttönen T, Hannonen P. Work disability in rheumatoid arthritis 10 years after the diagnosis. *J Rheumatol* 1999; 26:1681-5.
 106. Lubeck DP, Spitz PW, Fries JF, Wolfe F, Mitchell DM, Roth SH. A multicenter study of annual health service utilization and costs in rheumatoid arthritis. *Arthritis Rheum* 1986; 29:488-93.
 107. Wolfe F, Zwiilich SH. The long-term outcomes of rheumatoid arthritis: A 23-year prospective, longitudinal study of total joint replacement and its predictors in 1,600 patients with rheumatoid arthritis. *Arthritis Rheum* 1998; 41:1072-82.
 108. Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ, Spitz PW, Fries JF. The clinical value of the Stanford health assessment questionnaire functional disability Index in patients with rheumatoid arthritis. *J Rheumatol* 1988; 15:1480-8.
 109. Leigh JP, Fries JF. Mortality predictors among 263 patients with rheumatoid arthritis. *J Rheumatol* 1991; 18:1307-12.
 110. Callahan LF, Cordray DS, Wells G, Pincus T. Formal education and five-year mortality in rheumatoid arthritis: Mediation by helplessness scale scores. *Arthritis Care Res* 1996; 9:463-72.
 111. Callahan LF, Pincus T, Huston JW, III, Brooks RH, Nance EP, Jr., Kaye JJ. Measures of activity and damage in rheumatoid arthritis: Depiction of changes and prediction of mortality over five years. *Arthritis Care Res* 1997; 10:381-94.
 112. Söderlin MK, Nieminen P, Hakala M. Functional status predicts mortality in a community based rheumatoid arthritis population. *J Rheumatol* 1998; 25:1895-9.

113. Sokka T, Hakkinen A, Krishnan E, Hannonen P. Similar prediction of mortality by the health assessment questionnaire in patients with rheumatoid arthritis and the general population. *Ann Rheum Dis* 2004; 63:494-7.
114. Russak SM, Croft JD, Furst DE, Hohlbauch A, Liang MH, Moreland L et al. The use of rheumatoid arthritis health-related quality of life patient questionnaires in clinical practice: lessons learned. *Arthritis Care Res* 2003; 49:574-84.
115. Yazici Y, Sokka T, Ricciardi DD, Pincus T. Differences in ACR Core Data Set measures in different ethnic groups: implications for interpretation of clinical trial data. *Arthritis and Rheumatism* 50, S389. 2004. Ref Type: Abstract
116. Pincus T, Yazici Y, Bergman M, Goodin V, Sokka T, Swearingen C. Simplified formats of visual analog scales for use in standard rheumatology clinical care: agreement between a 10 cm horizontal line and 21 numbered circles. *Arthritis and Rheumatism* 54,S345-S346. 2006. Ref Type: Abstract
117. Pincus T. The American College of Rheumatology (ACR) core data set and derivative "patient only" indices to assess rheumatoid arthritis. *Clin Exp Rheumatol* 2005; 23:S109-S113.
118. Wells G, Li T, Maxwell L, Maclean R, Tugwell P. Sensitivity of measures of function and patient reported outcomes following treatment with abatacept in patients with rheumatoid arthritis. *Arthritis and Rheumatism* 52, S467. 2005. Ref Type: Abstract
119. Freiman JA, Chalmers TC, Smith H, Jr., Kuebler RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial: survey of 71 "negative" trials. *N Engl J Med* 1978; 299:690-4.
120. Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 1979; 301:1410-2.
121. Sackett DL. The competing objectives of randomized trials. *N Engl J Med* 1980; 303:1059-60.
122. Huskisson EC. Important factors in the success and failure of clinical trials (closing remarks). *Agents Actions Suppl* 1980; 7:323-4.
123. Freireich EJ. The randomized clinical trial as an obstacle to clinical research. In: Varco RL, Delaney JP, editors. *Controversy in Surgery*. 2 ed. Philadelphia: Saunders,W.B.; 1983. 5-12.
124. Feinstein AR. An additional basic science for clinical medicine: II. The limitations of randomized trials. *Ann Intern Med* 1983; 99:544-50.
125. Diamond GA, Forrester JS. Clinical trials and statistical verdicts: probable grounds for appeal. *Ann Intern Med* 1983; 98:385-94.
126. Chalmers TC, Celano P, Sacks HS, Smith H, Jr. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983; 309:1358-61.
127. Bombardier C, Tugwell P. Controversies in the analysis of longterm clinical trials of slow acting drugs (editorial). *J Rheumatol* 1985; 12:403-5.
128. Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S. Determining optimal therapy - randomized trials in individual patients. *N Engl J Med* 1986; 314:889-92.
129. Gotzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Controlled Clin Trials* 1989; 10:31-56.
130. Felson DT, Anderson JJ, Meenan RF. Time for changes in the design, analysis, and reporting of rheumatoid arthritis clinical trials. *Arthritis Rheum* 1990; 33:140-9.
131. Rothwell PM. Can overall results of clinical trials be applied to all patients? *Lancet* 1995; 345:1616-9.
132. Pincus T. Rheumatoid arthritis: disappointing long-term outcomes despite successful short-term clinical trials. *J Clin Epidemiol* 1988; 41:1037-41.
133. Pincus T, Wolfe F. Response to letter: Gold therapy for rheumatoid arthritis: Challenges to traditional paradigms. *Ann Intern Med* 1992; 117:169-70.
134. Pincus T. Limitations of randomized controlled clinical trials to recognize possible advantages of combination therapies in rheumatic diseases. *Semin Arthritis Rheum(Suppl 1)* 1993; 23:2-10.
135. Pincus T, Stein M. What is the best source of useful data on the treatment of rheumatoid arthritis: Clinical trials, clinical observations, or clinical protocols? *J Rheumatol* 1995; 22:1611-7.
136. Pincus T, Sokka T. Evidence-based practice and practice-based evidence. *Nat Clin Pract Rheumatol* 2006; 2(3):114-5.
137. Sokka T, Pincus T. Improved outcomes in consecutive patients with rheumatoid arthritis in 2000 compared to 1985, according to an identical cross sectional quantitative standard protocol to evaluate rheumatoid arthritis (SPERA). *Arthritis and Rheumatism* 46(9), S242. 2002. Ref Type: Abstract

138. Sokka T, Pincus T. Radiographic Larsen scores in patients with rheumatoid arthritis (RA) seen for standard care at the same site were substantially better in 2000 versus 1985, prior to availability of anti-tumor necrosis factor biological agents. *Arthritis and Rheumatism* 50, S458. 2004. Ref Type: Abstract
139. Pincus T, Sokka T, Kautiainen H. Patients seen for standard rheumatoid arthritis care have significantly better articular, radiographic, laboratory, and functional status in 2000 than in 1985. *Arthritis Rheum* 2005; 52:1009-19.
140. Pincus T, Sokka T, Chung CP, Cawkwell G. Declines in number of tender and swollen joints in patients with rheumatoid arthritis seen in standard care in 1985 versus 2001: possible considerations for revision of inclusion criteria for clinical trials. *Ann Rheum Dis* 2006; 65(7):878-83.
141. Sokka T, Pincus T. Eligibility of patients in routine care for major clinical trials of anti-tumor necrosis factor alpha agents in rheumatoid arthritis. *Arthritis Rheum* 2003; 48(2):313-8.
142. Sokka T, Kautiainen H, Horslev-Petersen K, Luukkainen R, Combe B, Sibilia J et al. Association of rheumatoid arthritis disease activity and gross domestic product in 20 countries in the QUEST-RA study. *Ann Rheum Dis* 66 (Suppl)[suppl 11], 274. 2007. Ref Type: Abstract
143. Austin HA, III, Klippel JH, Balow JE, le Riche NGH, Steinberg AD, Plotz PH et al. Therapy of lupus nephritis: controlled trial of prednisone and cytotoxic drugs. *N Engl J Med* 1986; 314:614-9.
144. Felson DT, Anderson JJ, Meenan RF. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis: results of two metaanalyses. *Arthritis Rheum* 1990; 33:1449-61.
145. Pincus T, Marcum SB, Callahan LF. Long-term drug therapy for rheumatoid arthritis in seven rheumatology private practices: II. Second-line drugs and prednisone. *J Rheumatol* 1992; 19:1885-94.
146. Felson DT, Anderson J. Evidence for the superiority of immunosuppressive drugs and prednisone over prednisone alone in lupus nephritis: results of a pooled analysis. *N Engl J Med* 1984; 311:1528-33.
147. Ruberman W, Weinblatt E, Goldberg JD, Chaudhary BS. Psychosocial influences on mortality after myocardial infarction. *N Engl J Med* 1984; 311:552-9.
148. Weinblatt E, Ruberman W, Goldberg JD, Frank CW, Shapiro S, Chaudhary BS. Relation of education to sudden death after myocardial infarction. *N Engl J Med* 1978; 299:60-5.
149. Yazici Y, Erkan D, Paget SA. Monitoring methotrexate hepatic toxicity in rheumatoid arthritis: is it time to update the guidelines? *J Rheumatol* 2002; 29:1586-9.
150. Yazici Y, Erkan D, Harrison MJ, Nikolov NP, Paget SA. Methotrexate use in rheumatoid arthritis is associated with few clinically significant liver function test abnormalities. *Clin Exp Rheumatol* 2005; 23(4):517-20.
151. Pincus T, Koch GG, Sokka T, Lefkowitz J, Wolfe F, Jordan JM et al. A randomized, double-blind, crossover clinical trial of diclofenac plus misoprostol versus acetaminophen in patients with osteoarthritis of the hip or knee. *Arthritis Rheum* 2001; 44:1587-98.
152. Pincus T, Koch G, Lei H, Mangal B, Sokka T, Moskowitz R et al. Patient preference for placebo, acetaminophen or celecoxib efficacy studies (PACES): two randomized placebo-controlled cross-over clinical trials in patients with osteoarthritis of the knee or hip. *Ann Rheum Dis* 2004; 63:931-9.
153. Spiro HM. Doctors, patients, and placebos. New Haven: Yale University Press; 1986.
154. Ader R. The placebo effect as a conditioned response. In: Ader R, Weiner H, Baum A, editors. *Experimental Foundations of Behavioral Medicine: Conditioning Approaches*. Hillsdale, NJ: Erlbaum; 1988. 47-66.
155. Chaput de Saintonge DM, Herxheimer A. Harnessing placebo effects in health care. *Lancet* 1994; 344:995-8.
156. Ernst E, Herxheimer A. The power of placebo: Let's use it to help as much as possible. *Br Med J* 1996; 313:1569-70.
157. Kaptchuk TJ. Powerful placebo: the dark side of the randomised controlled trial. *The Lancet* 1998; 351:1722-5.
158. Ratain JS, Hochberg MC. Clinical trials: a guide to understanding methodology and interpreting results. *Arthritis Rheum* 1990; 33:131-9.
159. Pincus T, Sokka T. Should contemporary rheumatoid arthritis clinical trials be more like standard patient care and vice versa? *Ann Rheum Dis* 2004; 63 ((Suppl II)):ii32-ii39.
160. Pincus T, Yazici Y, Bergman M, Swearingen C, Harrington T. A proposed approach to recognise "near-remission" quantitatively without formal joint counts or laboratory tests: a patient self-report questionnaire routine assessment of patient index

data (RAPID) score as a guide to a "continuous quality improvement" strategy. Clin Exp Rheumatol 2006; 24 (6 Suppl 43):S-60-5; quiz S-66-73.

161. Pincus T, Yazici Y, Bergman M. A practical guide to scoring a Multi-Dimensional Health Assessment Questionnaire (MDHAQ) and Routine Assessment of Patient Index Data (RAPID) scores in 10-20 seconds for use in standard clinical care, without rulers, calculators, websites or computers. Best Pract Res Clin Rheumatol 2007; 21(4):755-87.
162. Pincus T, Yazici Y, Bergman M, Maclean R, Harrington T. A proposed continuous quality improvement approach to assessment and management of patients with rheumatoid arthritis without formal joint counts, based on quantitative Routine Assessment of Patient Index Data (RAPID) scores on a Multidimensional Health Assessment Questionnaire (MDHAQ). Best Pract Res Clin Rheumatol 2007; 21(4):789-804.
163. Pincus T, Sokka T. Can a Multi-Dimensional Health Assessment Questionnaire (MDHAQ) and Routine Assessment of Patient Index Data (RAPID) scores be informative in patients with all rheumatic diseases? Best Pract Res Clin Rheumatol 2007; 21(4):733-53.
164. Pincus T, Bergman MJ, Yazici Y, Hines P, Raghupathi K, Maclean R. An index of only patient-reported outcome measures, routine assessment of patient index data 3 (RAPID3), in two abatacept clinical trials: similar results to disease activity score (DAS28) and other RAPID indices that include physician-reported measures. Rheumatology (Oxford) 2008; 47(3):345-9.
165. Yazici Y, Bergman M, Pincus T. Time to score various measures to assess rheumatoid arthritis. J Rheumatol. In press 2008.
166. Pincus T, Bergman M, Sokka T, Roth J, Swearingen C, Yazici Y. Visual analog scales in formats other than a 10 cm horizontal line to assess pain and other clinical data. J Rheumatol. In press 2008.
167. Callahan LF, Brooks RH, Summey JA, Pincus T. Quantitative pain assessment for routine care of rheumatoid arthritis patients, using a pain scale based on activities of daily living and a visual analogue pain scale. Arthritis Rheum 1987; 30:630-6.
168. Callahan LF, Pincus T. Formal education level as a significant marker of clinical status in rheumatoid arthritis. Arthritis Rheum 1988; 31:1346-57.

ŞEKİL AÇIKLAMALARI

İSKANDINAV VERİ TABANLARI

Tuulikki Sokka

İskandinavya, bildiğiniz gibi, kuzey kutbuna yakın neredeyse dünyanın tepesinde bir bölgedir. Danimarka, Finlandiya ve Norveç'in her birinde 5 milyon civarında, İsveç'te 9 milyon civarı ve İzlanda'da yarım milyondan az sayıda olmak üzere toplam yaklaşık 24 milyon kişi yaşamaktadır. Gayrisafi milli hasıla makul seviyededir ve yaşam beklentisi oldukça yüksektir. Popülasyon kayıt sistemleri kuzey ülkelerine hastır ve yaklaşık 16. yy kadar eskiye dayanmaktadır. İzlanda'da ise popülasyon veri tabanı 1100 yıl kadar eskiye gider.

Kuzey ülkeleri arasında veri tabanları bakımından bazı küçük farklar olduğu kadar benzerlikler de vardır. En önemlisi, veri tabanlarının her ülkede bir diğeriyle bağlantılı olmasıdır ve bir kişi tek bir kimlik koduyla tüm veri tabanlarında bulunabilir. Bu veri tabanlarının özelliklerini size aktaracağım.

Finlandiya'daki popülasyon veri tabanı, ülkede sürekli yaşayan tüm vatandaşlar ve yabancıları kapsar. Veri tabanında kişisel kimlik kodu, kişinin adı, soyadı, cinsiyeti, şimdiki ve önceki adresi, medikal durumu, eş, çocuklar, ebeveyn bilgileri ve ölüm tarihi yer alır.

Kanser veri tabanı, çoğu İskandinav ülkesinde 1950'li yıllarda kurulmuştur, 1960'ta zorunlu hale gelmiştir. Tüm kanser tipleri ile ilgili verileri toplar. Diş hekimleri, hastaneler ve tıbbi laboratuvarlar bu veri tabanına veri gönderir.

Ölüm nedeni veri tabanı, ICD kodlarına göre 1900'lü yılların başından bu yana ölüm tarihleri ve nedenlerini kaydeder.

Hastanede yatan hasta veri tabanı, yatan hastalarla ilgili bilgileri içerir ve ICD kodlarına göre tanıyı, uygulanan prosedürleri, hastane kodlarını, yatış tarihini ve taburcu olma tarihini kaydeder. Veriler devlet tarafından yönetilen veri tabanlarında bulunur ve diğer veri tabanlarıyla bağlantılıdır.

İlaç geri ödeme veri tabanı: Finlandiya'da sosyal güvenlik kurumu belli kronik hastalıklar için geri ödeme yapar ve aynı zamanda kişisel kimlik kodu, hastalığın tarihi, hastalık grubunun geri ödeme kodunu da içerir. Bu veri tabanı bölgedeki (veya ülkedeki) belli kronik koşullarla tanı konan tüm hastaları bulmak için kullanılabilir.

*Jyväskylä Merley Hastanesi,
Jyväskylä, Finlandiya*

Reçete veri tabanı, reçete edilen ilaçlarla ilgili verileri içerir. Örneğin bir hasta romatoid artrit için geri ödenen bir ilaç alıyorsa, hastanın gerçekten de bu ilaçları satın alıp almadığı kontrol edilebilir.

İsveç'te çoğul cinsiyet veri tabanı vardır. En heyecan verici veri tabanı İzlanda'daki gen veri tabanıdır. Bu veri tabanı İzlanda'da yaşamış olan herkesi içerir ve dahası çoğu İzlandalı'nın genotipini de içerir.

Kuzey biyolojik örnek bankası, veri tabanlarına bağlantılıdır. Bu biyobanka, İskandinav ülkelerindeki kanser vakalarının nedenleri ve kontrolü için kurulmuştur, prematürite programları, popülasyon çalışmaları ve diğer kaynaklardan gelen kan numunelerini toplar.

Neden bu kadar çok veri tabanı var? Buna % 100 bir cevap verebileceğimizi sanmıyorum ama en azından en önemli sebebinin vergilerin daha etkin şekilde toplanmasını sağlamak ve sosyal anlamda bilgileri arşivlemek olduğunu söyleyebilirim. Kanunla yürürlükte ama kanunî bir kısıtlama yoktur. Ancak, kanunlar bu veri tabanlarının sadece veri toplamaktan öte kurumlar yöneticiler tarafından kullanılmasını da destekler. Veri tabanları sadece veri toplamaz, en önemli konu verinin kişilere, kurumlara ve otoritelerle diğer veriye ihtiyaç duyan olaylara hizmet etmesidir. Güvenilirlik ve kalite çok önemlidir ve verinin güvenilirliğinin sağlanması için güvenilirlik çalışmaları yürütülmektedir.

Popülasyon bilgileri hem halk hem de özel servisler amacıyla toplanmaktadır. Sivil otoriteler, gerekli bilgileri vatandaşlardan değil doğrudan sistemden alır; Ayrıca Finlandiya'da kanunlar, bu bilgilerin bilimsel amaçla kullanımına da izin verirler.

Gizlilik önemlidir. Her vatandaşın kendi bilgilerini belli amaçlar için; örneğin pazarlama, adres aranması, kilise veya gen katalogları gibi konularda yasaklama hakkı vardır. Ancak çok az insan bu haktan yararlanır. Sağlıkla ilgili veri tabanları çoğu ülkede önemli bir sorundur ama İskandinav'da veritabanları toplum düzeninin bir parçası olarak kabul edilir. Aslında, sağlık bilgileri ile ilgilenen veri tabanları kişileri kontrol etmek ve yönetmek için geliştirilmemiştir. Amaç

halk sağlığını iyileştirmek ve halk sağlığı stratejileri geliştirmek için halk sağlığını tanımdır. İskandinav'da toplum, hasta üyelerinden sorumludur ve ücretsiz veya düşük maliyetli tıbbi bakım ile vatandaşlarına rehabilitasyon aylığı ve yararları sağlar.

Rheumatic Disease Clinics in North America dergisine İskandinav hastalar üzerine bir makale yazarken, tüm bunları insanların anlayacağı bir dilde ifade etmeye çalıştım, çünkü gizlilik düzenlemeleri ve kişisel sağlık bilgilerinin korunması çoğu ülkede hasta veri tabanlarını engellemektedir. Toplumun hasta üyesinden sorumlu olması ile ilgili tarihsel geçmişi açıklamak isterim. Bazı ülkelerdeki kanunlar bu tür veri tabanlarını birey haklarına aykırı bulur. Halbuki İskandinav'da bireyin sağlığı ön plandadır. İskandinav kanunları gizlilikten ziyade veritabanlarının eksiksiz olmasını gözetir.

Bu veri tabanları, medikal araştırmalarda, örneğin romatoid artrit hastaların mortalitesi, romatolojik hastalıkların insidansı ve kanser, romatolojik maliyetler ve maluliyet ile ilgili konulardaki romatolojik araştırmalarda çok kullanılmıştır. Klinik medikal araştırma ile ilgili pek çok araştırma programları bu veri tabanlarından yararlanır. Veri tabanlarını normal bir prosedür olarak diğer ülkelerde pek olmayan çok yararlı bir araç olduğunu pek farketmeden kullanırız.

Bir önceki gün İsveç'te veri tabanlarının etkin bir biçimde nasıl kullanıldığına ilişkin güzel örnekleri dinledik. Şimdi Orta Finlandiya'da kendi çalışmalarında veri tabanlarının kullanımı üzerine bir kaç örnek vereceğim.

Orta Finlandiya romatoid artrit (RA) veri tabanı, kliniğimizde 1980'lerin ortasından bu yana görülen RA'li hastaları içerir; Haziran 2003'te 1500 kişi veri tabanına eklenmiştir. Hastalarımızı popülasyona göre karşılaştırabileceğimiz bir popülasyon örnekleme oluşturmak istedik. Bu nedenle, popülasyon veri tabanı merkezine başvurduk ve cinsiyet ve yaşla uyumlu örnekleme aldık. RA'li hastalar ve popülasyon örnekleme fonksiyonel kapasite, ağrı, egzersiz alışkanlıkları, sigara içme durumu ve pek çok başka soruları içeren bir anketi posta ile gönderdik. Hastaların fonksiyonel kapasitelerini popülasyon örnekleme ile karşılaştırdık.

İki sene sonra, benzer bir anketi posta ile gönderdik ve ölüm veri tabanından ölümleri kontrol ettik. Azalmış fonksiyonel kapasitenin 2 yıl içinde romatoid artritli hastalarda popülasyona benzer mortalite öngördüğünü belirledik .

Mevcut projelerde, RA prevalansının tahmin edilmesi için geri ödeme veri tabanından romatizmal hastalıklar için geri ödeme alan hastalarla ilgili veri istedik. Dahası, uzun süreli takip için yeni bir popülasyon örnekleme talep ettik.

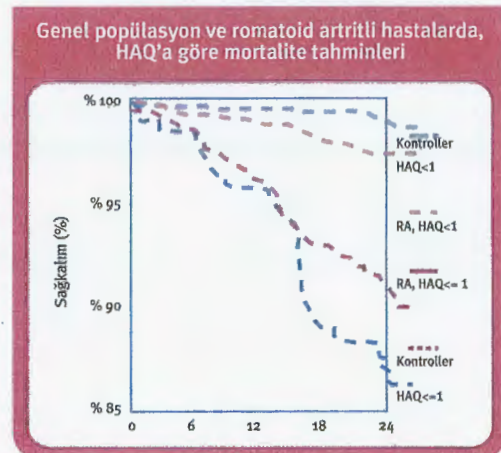
Bu örnekler, umarım ki ulusal popülasyon veri tabanlarının İskandinav'daki epidemiyoloji araştırması için mükemmel bilgi sağladığını göstermiştir. Bu mükemmel veri kaynakları daha sık kullanılmalıdır.

Dr. Yazıcı: Popülasyon veritabanı ve RA veri tabanının güncellemesi ile ilgili biraz daha bilgi verir misiniz?

Dr. Sokka: Popülasyon veri tabanından başlayalım. Örneğin bir bebek sahibi olduğunuzda her şey başlar. Bebek doğar ve hemen hastaneden popülasyon veri tabanına yenidoğanla ilgili bir not gider. Bir süre sonra bebek için sosyal güvenlik numarası verilir. Veri tabanında bulunan bebekle ilgili bilgiler ebeveyn isimleri, sosyal güvenlik numaraları, adres vs. bilgileri içerir. Öldüğünüz zaman da bu veriler aynı şekilde işlenir. Hangi kurum sizinle ilgili bilgileri işliyorsa, ölüm kayıtları oraya gönderilir. Her taşınma durumunda kişiler, veri tabanına bilgi vermekle yükümlüdür. Veri tabanları, verilerin doğruluğu bakımından kontrol edilir. Romatizmal hastalıklarla ilgili çok sayıda veri tabanı vardır ve bir tanesi Orta Finlandiya'da bulunur. Orta Finlandiya RA veri tabanına temel olarak 1980'den bu yana romatoid artritli hastalar dahil edilmiştir. Önce medikal kayıtlarda retrospektif verileri topladık. Hastanemiz, Orta Finlandiya'da RA hastalarının bakıldığı tek yerdir ve çeyrek milyon hastayı kapsar. Ocak 1997'den bu yana tüm yeni hastalar veri tabanına prospektif olarak dahil edilmiştir ve veriler 2 yıl boyunca her ziyarette güncellenir. Hastalar 5 ila 10. yıllarda takip ziyaretleriyle izlenir. Tam zaman çalışan bir eleman ile bu veri tabanının güncellenmesi ancak



Ref: Sokka T, et al. Arthritis Rheum 2003; 48:59-63.



mümkün olduğundan, bu sistem çok zaman alıcıdır; bu nedenle herkese bu tarz bir RA veri tabanı kullanmasını öneremem. Şimdilerde, otomatik veri girişli daha modern bir sistemin hazırlanması için çalışmaktayız.

Dr. Yazıcı: Sorunun üçüncü ve son kısmını sorabilir miyim? Kohortunuzun yaşam beklentisini serbest popülasyondaki beklentiyle nasıl kıyaslıyorsunuz?

Dr. Sokka: Bu istatistikçinin sorunudur.

TÜRKİYE ANTI-TNF VERİ TABANI

Salih Pay

*Gülhane Askeri Tıp Akademisi,
Ankara*

İkibinli yılların başında Türk romatologları TNF antagonistleri ile tanıştı. İlk infliksimab ruhsat aldı. Bunu etanersept ve adalimumab izledi. Çok etkili bulduğumuz ve sık kullanmaya başladığımız bu ilaçlar yanlarında sorunlar da getirmeye başladı. İlk önemli sorun sanırım tüberkülozdu. Romatoloji Araştırma ve Eğitim Derneği bu sorun ile ilgili olarak 2002 ve 2005 yıllarında iki toplantı yapmış ve iki kılavuz yayımlamıştır. Bu kılavuzlar sayesinde tüberküloz ile ilgili olarak bu ilaçları daha güvenli uygulamaktayız.

Son yıllarda TNF antagonistleri ile malignite arasında bir ilişki olup olmadığı konusunda endişeler ortaya çıkmıştır.

2006 yılının Haziran ayında RAED yönetim kurulu olarak Türkiye’de bizim üyelerimizin bu ilaçları kullandıkları hastalarda TNF antagonistleri ile malignite gelişimi arasında ilişki olup olmadığını araştırmaya karar verdik. Bu araştırmanın ilk sonuçlarını Mart ayında Kartepe’de yapılan RAED genel kurul toplantısında tartışmıştık. Ben bu sonuçları sizlere tekrar aktaracağım. Aslında romatoloji pratiğinde kansere çok uzak değiliz. Romatolojik hastalıkların ayırıcı tanısında paraneoplastik sendromların yer aldığını çok iyi biliyoruz. Polimiyozit bunların en iyi örneğidir. Polimiyozitin dışında bir çok romatolojik klinik tabloya benzer bulgular çeşitli kanserlerde de görülmektedir. Romatoloji uzmanının kanser ile ilgili diğer bir sorunu da hastaların takibinde ortaya çıkmaktadır. Biz çok iyi biliyoruz ki özellikle otoimmün kökenli hastalıklar takip esnasında lenfoproliferatif maligniteler yönünden risk taşımaktadır. Bu açıdan bakıldığında topluma göre 44 kat malignite risk artışı gösteren Sjögren sendromu en riskli gibi görünmektedir. Romatoid artritli hastalarda ise en az 5 çalışmada malignite riskinin 2 ile 4 kat arasında arttığı gösterilmiştir. Bu risk artışının yüksek klinik aktivite ile ilişkili olduğu saptanmıştır. Sistemik lupus eritematozuslu hastalarda malignite riskinin araştırılan çalışmalarda çelişkili sonuçlar alınmıştır. Bazı çalışmalar malignite riskinin arttığını gösterirken tersini iddia edenler de vardır. Sklerodermada ise solid tümör olan akciğer kanserinde artış olduğu bildirilmiştir. Bunun yanı sıra son yıllarda bazı tümör markerlarının

romatoid artrit gibi inflamatuvar hastalıklarda arttığı ve klinik aktivite ile korelasyon gösterdiği bildirilmiştir. Ayrıca romatolojik hastalıkların tedavisinde kullandığımız bazı sitotoksik ilaçların da maligniteye neden olduğu iyi bilinmektedir.

TNF antagonistlerinin de sitotoksik ilaçlar gibi malignite riski taşıyıp taşımadığı tam olarak bilinmemektedir. Bizim çalışmamız pilot bir çalışmadır. Romatolojik hastalığı olan Türk hastaların TNF antagonisti kullanımı ile neoplazi gelişimi olasılığı arasındaki ilişkiyi ortaya koymak için yapılmıştır. Bu verileri toplarken oluşturduğumuz formu görüyorsunuz. Bu form katılan merkezlerin isimleri, hastaların doğduğu yerler, yaşadığı yerler, tanıları, hastalık sürelerini içeriyor. Bu formda hastanın kullandığı TNF antagonisti veya antagonistlerinin dozu, kullanım süresi de yer almaktadır. Ayrıca formda hastada mevcut bir premalign lezyon olup olmadığı da sorgulanmıştır. Tedavi esnasında malign hastalık tespit edildiye, tanısı, tanı tarihi ve prognozu da formlara kaydedildi.

Verilerin toplanması ile ilgili biri olumlu birisi olumsuz iki önemli konuyu belirtmem gerekiyor. Olumlu olan RAED üyeleri, daha çok eğitim hastanelerde ve üniversitelerde çalışmaktadır. Bu nedenle verilerin düzenli tutulduğunu ve güvenli olduğunu söylemek mümkündür. Olumsuz olan ise tam tersi olarak verilerin toplanmasında kontrollerin olmaması bu çalışmanın önemli bir eksiğidir.

Biz bu çalışmada bulgularımızı Fidaner ve arkadaşlarının 1993 – 1994 yıllarında İzmir ve çevresinde yaptıkları kanser kayıt sistemini temel alarak değerlendirdik. Bu çalışmada toplam malignite oranları erkeklerde yüzbinde 157,5 kadınlarda ise yüzbinde 94 olarak bildirilmektedir. Ayrıca bu çalışmada maligniteler yaş, cins ve malignite türüne göre de hesaplanmıştır.

Bizim çalışmamıza 26 merkezimiz katıldı ve 2395 hastanın verisine ulaştık. Bu hastaların toplam kullandıkları TNF antagonistleri süresi 2555 hasta yılıydı. Toplam 15 adet malignite tespit ettik. Bu malignitelerin tanısal dağılımına baktığımızda biri non Hodgkin lenfoma ve diğeri Hodgkin lenfoma olmak üzere toplam iki adet

lenfoproliferatif tümör ve 13'ü ise çeşitli tanılarda solid tümörlerdi. Solid tümörlerin dağılımında ise 5 adet kolon kanseri, 2 adet meme kanseri ve diğer toplam 6 adet farklı türlerde tümörler vardı. Bu malignitelerin ilaçlara göre dağılımına baktığımızda 10 malignitenin etanersept, 3 malignitenin infliksimab ve 1 malignitenin ise adalimumab kullanan hastalarda ortaya çıktığı görülmektedir. Son 1 malignite ise önce etanersept ardından adalimumab kullanan hastada saptanmıştır.

Çalışmaya katılan 282 hasta total 224 hasta yılı adalimumab, 1011 hasta 1029 hasta yılı etanersept ve 925 hasta 1015 hasta yılı infliksimab kullanmıştı. Yüzyetmişyedi hastanın ise farklı kombinasyonlarda bu üç ilacı ardışık kullandığını tespit ettik. Tedavide TNF antagonistlerinin başlıca değiştirme nedeni etkinliğinin yetersiz olması veya yan etkilerin ortaya çıkmasıydı.

Bu verilerimizi Fidaner ve arkadaşlarının İzmir ve çevresinde yaptıkları çalışma ile karşılaştırdığımızda SIR değerlerinin total hastalarda 1.26, etanersept alanlarda 2.30, adalimumab alanlarda 0.62 ve infliksimab alanlarda ise 0.66 olduğunu gördük. Bu değerler istatistik olarak anlamlı değildi. Vaka gruplarının verileri geniş bir güven aralığına sahip ve güven aralıkları üst üste binmiş durumdaydı.

Çalışmada saptadığımız malignitelerin kadın-erkek dağılımına baktığımızda total 15 malignitenin 8 tanesinin kadın 7 tanesinin erkek hastalarda ortaya çıktığını gördük. Kadın hastalarımız erkek hastalardan biraz daha fazlaydı. Toplam 1367 kadın hasta 1520 hasta yılı TNF antagonisti kullanırken, 1028 erkek hastamız ise 1034 hasta yılı bu ilaçları kullanmıştı. Ancak SIR değerlerinin erkeklerde 1.66 ve kadınlarda 1.68 olduğunu saptadık.

Çalışmada saptanan malignitelerin TNF antagonistlerinin kullanım sürelerine göre değerlendirdiğimizde birinci yılda 8, ikinci yılda 5, üçüncü ve dördüncü yıllarda ise birer adet malignite tespit ettik. Bu verilere göre kümelenmenin bir ve ikinci yıllarda olduğu görülmektedir. Ancak şu gerçeği unutmamamız da gerekir ki bizim total 2395 hastamızın hepsi 4 yıl süresince ilaç kullanmadı. Yani birinci yılın sonuna ulaşan hasta sayısı ile ikinci yılın sonuna

ulaşan hasta sayısı aynı değil. Yaptığımız hesaplamalarda TNF antagonisti kullanan hastalarda malignite ortaya çıkma riski birinci yılda % 0.48, ikinci yılda % 0.68 ve üçüncü yılda ise % 0.43 oranlarındaydı ve bu değerler arasında istatistik fark yoktu.

Çalışmamızda saptanan malignitelerin yaş gruplarına göre dağılımlarına baktığımızda üçüncü ve dördüncü dekadlarda saptanan toplam iki malignite dışında daha çok beş ve daha üzeri dekadlarda ortaya çıktığı görülmektedir. Bu dağılımın kadın/erkek ayrımı yapıldığında da değişmediği görülmektedir. Çalışmamızda saptanan tümörlerin yaş gruplarına göre dağılımlarını İzmir çalışması ile kıyasladığımızda yaşa spesifik SIR değerlerinin oldukça örtüştüğünü ve birbirinden farksız olduğunu saptadık. Ancak etanersept kullanan hasta grubunun ayrı olarak değerlendirdiğimizde 15-24, 45-54 ve 65 yaş üzeri gruplarda yaş spesifik malignite oranlarının etanersept kullanan hasta grubunda biraz yüksek olduğu bulunmuştur.

Çalışmamızda saptanan maligniteleri hastaların tanılarına göre de sınıflandırdık. 15 malignitenin 11'i romatoid artritli, 2'si ankilozan spondilitli ve 2'si ise psöriyatik artritli hastalarda ortaya çıktı. Çalışmaya katılan 1006 ankilozan spondilitli hasta 1027 hasta yılı, 1036 romatoid artritli hasta 1145 hasta yılı ve 144 psöriyatik artritli hasta 141 hasta yılı TNF antagonisti kullanmıştı. Romatoid artritli hastalarda ankilozan spondilitli hastalara göre daha fazla malignite saptanmasına karşın bu hasta grupları arasında hasta sayısı ve TNF antagonisti kullanım süresi açısından fark yoktu. Ancak hastaların demografik özellikleri oldukça farklıydı. Örneğin romatoid artritli hastaların % 83.4'ü kadın, % 16.6'sı erkek ve yaş ortalamaları 49.1±12.7 idi. Ankilozan spondilitli hastaların ise % 30.5'i kadın, % 69.5'i erkek ve yaş ortalaması ise 37.3±10.7 olarak saptandı. Küçük bir grup olan psöriyatik artritli hastaların ise % 54.2'si kadın, % 55.8'i erkek ve yaş ortalaması 43.6±12.5 olarak tespit edildi. Saptanan malignitelerin klinik tanılarına göre SIR değerleri ankilozan spondilitli hastalarda 0.66, romatoid artritli hastalarda 1.36, psöriyatik artritli hastalarda ise 3.83 olarak saptandı. Tüm hastalarda saptanan SIR değerlerinin güven

aralıklarının alt sınırlarının 1'in altında olduğu görülmektedir. Bu haliyle TNF antagonistlerinin hastalık gruplarında malignite artışına neden olmadığını söylemek mümkündür. Psöriyatik artritli hastalarda saptanan abartılı SIR değeri olasılıkla olgu sayısının çok düşük olmasından kaynaklanmaktadır. Son olarak romatoid artritli hastalarda kullandığımız ilaçlar arasında fark var mı diye baktığımızda, SIR değerinin tüm ilaçlar için 1.36, etanersept için 2.46, adalimumab için 0.64 ve infliksimab için 0.75 olduğunu saptadık.

Göreceli olarak küçük bir kohorttaki bu pilot sonuçlar bazı ileri çalışmaların yapılması gerektiğini göstermektedir: Birincisi TNF antagonisti kullanımı ile ilgili farklı hastalıklarla farklı neoplazilerin birlikteliği var mı? İkincisi farklı anti-TNF ajanlar farklı neoplaziler ile ilişkili midir?

Soru: Bir şeyi açabilir miyiz? Türkiye'deki kanser 93 – 94 yılları arasındaki veri ile karşılaştırıyorsunuz. Bu eleştiri alabilir mi?

Cevap: Kesinlikle eleştiri alır. 1993 ile 2007 arasında kanser tarama metodları çok gelişti. Türkiye'de verilerimizi karşılaştıracığımız başka veri yok.

Soru: Birkaç noktayla ilgili sorum olacak. Anti-TNF'lerle görülen tüberküloz konusuna değindiniz. Anti-TNF tedavisine başlamadan önce her birey için röntgen istiyor musunuz? Bu durumda bir şekilde akciğer kanseri dahil herhangi bir akciğer problemini taramaktan oluşan bir bğıntı kurmuş oluyorsunuz. Tabii ki akciğer kanserini hariç tutup diğer tüm kanser türlerine de bakacak olursanız bu sorun yaratır. Buna katılıyor musunuz?

Cevap: Evet, ancak tüberküloz yönünden bu araştırmayı yapmak zorundayız.

QUEST-RA: GLOBAL BİR VERİ TABANI

Tuulikki Sokka

Bir kısmınızın da katılmış olduğu QUEST – RA veri tabanından bahsetmek istiyorum. Amacımız en az 10 ülkede ülke başına 3 merkez olacak şekilde RA'li 100 hasta olarak kesitsel bir çalışma oluşturmaktır. Yararlı bulduğumuz monitorizasyon araçlarını romatologlara sunmak istedik: 4 sayfalık bir hasta anketi ve 3 sayfalık klinik değerlendirme. Bu durum ayrıntılı planlamayı gerektirdi; programa katılabilecek tüm arkadaşlara ve kişilere haber verdik. Hasta anketlerini çeşitli dillere tercüme ettik.

Benim kliniğim bir örnektir. Hastalar bizi görmeye gelmeden önce anketlerini doldururlar. Biz de belli protokole göre kaydederiz. Hastanın kendi kendine anketi doldurmasının ve sadece gerektiğinde hemşireden yardım almasının çok önemli olduğunu hatırlatmak isterim. Bu biraz şansla ilgili bir konudur, tüm merkezlerde bu kadar iyi olup olmadığından emin değilim. Çalışma arkadaşlarımızın klinik günlük rutin bakım sırasında veri toplamanın yararlarını görmelerini isteriz.

Hasta anketinde orijinal Sağlık Değerlendirme Anketi (HAQ) çok boyutlu HAQ'tan ilave sorular ve İtalya'dan ROAD soruları dahil genişletilmiş bir HAQ bulunur. Anket hastanın kendi kendine eklem sayımı, sigara içme ve fiziksel egzersiz gibi yaşam tarzı seçimlerini içermektedir.

Veri toplamaya 2005 yılında başladık ve geçen ay itibarıyla 21 ülkede, 58 merkezden neredeyse 5200 hastamız vardı. Bence bu olağanüstüdür. Herkes birbiriyle işbirliği içindeydi ve sadece posta, baskı gibi masraflar ödeniyordu. Bu program katılımcılarını, abstraktların ve yayınların yazılması süreci dahil olmak üzere, çok heyecanlandırdı.

Bu slaytta çoğu Avrupa'dan olan ülkeler gösterilmektedir. En yakında Kanada ve Rusya da katıldı. Türkiye, 2005 yılında programı ilk tamamlayan ülkelerdendir; Finlandiya, İsveç ve Danimarka da öyledir.

5200 hastanın verileri, veri tabanına dahil edildi. Hasta popülasyonunun RA'li hastalarda diğer veri setlerine benzer olması sürpriz değildir: çoğu kadın (% 78), ortalama yaşı 57, eğitimi 11 yıl, çoğu beyaz, % 73'ten fazlası romatoid faktör için pozitif bir hasta grubu.

*Jyväskylä Merley Hastanesi,
Jyväskylä, Finlandiya*

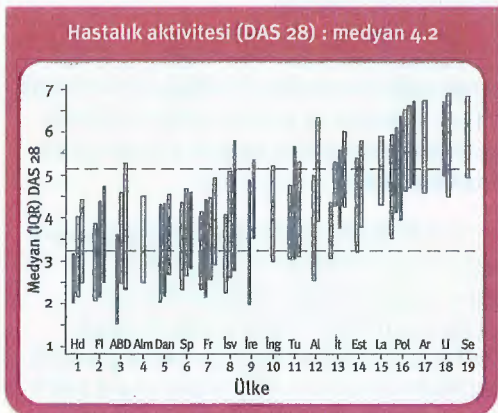
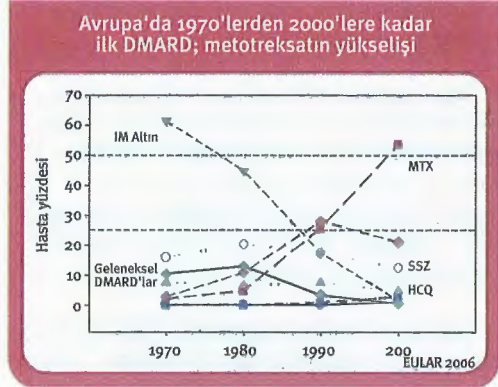
Nisan 2007'de QUEST - RA Ülkeleri



4300 hastanın analizine dayanan sonuçları paylaşmak istiyorum. Ardından 4800 kişi daha analize dahil edilmiştir. Bazı sonuçlar 15 ülkenin, bazıları 20 ülkenin verilerine dayanır. İlk analizlerden biri, 9 Avrupa ülkesinde başlangıçtaki hastalık modifiye eden ilaçları (DMARDlar) ilgilendirir ve bu EULAR 2006'da sunulmuştur. RA için ilk DMARD, 1970'lerde intramüsküler altın preparatıydı ama şimdi Avrupa ülkelerinde RA için ilk DMARD olarak % 5'ten az hastada kullanılmaktadır. 1970'lerde metotreksat, ilk DMARD olarak RA için kullanılmıyordu. Şimdi oran hala düşük olmasına rağmen % 50'den fazla RA hastası için ilk ilaç olarak metotreksat kullanılmaktadır.

DMARD'ların kullanımında ülkeler arasında farklar vardır. Prednison hastaların % 26-% 88, metotreksat % 68-% 92, hidroksiklorokin % 15-% 74, sülfasalazin % 6-% 84 ve biyolojik ajanlar % 2-% 53'ünde kullanılmıştır. En ileri merkezlerin seçilmiş olması, Fransa gibi bazı ülkelerde biyolojiklerin yüksek oranda kullanılmasının nedeni olabilir. Dahası, klinik araştırmalara katılanları hariç tutmadık. Bununla birlikte RA için nasıl tedavi edildiğiniz, nerede yaşadığınızla bağlı gibidir.

Ondokuz ülkede hastalık aktivitesini DAS28 kullanarak analiz ettik. DAS28 hassas ve şiş eklemler için 28 eklem sayımını, sedimantasyon hızını ve 0-10 arası değişen hastanın global hastalık değerlendirme skalasını içermektedir. DAS28'de düşük hastalık aktivitesinin 3.2'nin altında olduğunu ve yüksek aktivitenin 5.1 veya daha fazla olduğunu biliyoruz. Bazı ülkelerde ve kliniklerde hastaların çoğu düşük hastalık aktivitesine sahiptir. Diğer bazı ülkelerde, çoğu hastanın hastalık aktivitesi yüksektir. Bu 3 Finlandiya merkezinde hastalık aktivitesi hep benzerdir ama yine birbirine çok benzer olma eğilimi gösteren 6 Polonya merkezinden farklıdır. Türkiye burada ortadadır ve Türkiye merkezleri birbirine benzerdir. DAS28 burada ortadaki küçük siyah çizgi ile medyan olarak verilmiştir ve kolonlar da % 25 ila 75 içindeki kişilerdir. Bu programa ilk başladığımız zaman, sadece Finlandiya, Almanya ve Hollanda gibi batı ülkelerini dahil etmek istemiştik. Şimdi bu çalışmaya katılmak üzere Estonya, Macaristan ve ilk başta hiç düşünmediğimiz ülkelerde çok büyük bir heyecan gözleyoruz. Farklı ülkeleri



6 Ülkede 8 merkezde hastaların çoğunda düşük hastalık aktivitesi DAS28s 3.2 görülmüştür. Başka 7 ülkede 11 merkezde ise hastalarda yüksek hastalık aktivitesi (DAS> 5.1) saptandı.

Ref: Sokka et al. ARD 2007

dahil ettiğimizde bu sefer RA hastalığı ile ilgili daha iyi bir resmimiz olacak.

Sonraki slayt DAS28, ağrı, hastanın global değerlendirmesi, 0-10'a normalleştirilmiş HAQ malüliyet indeksi ve 0-10 ölçeğinde hekimin doktor global değerlendirmesini gösterir. Ülkeler, DAS28'e göre en düşükten en yükseğe doğru verilmiştir. Hasta bazı ölçülerin ikisi de aynı fenomeni ölçmektedir.

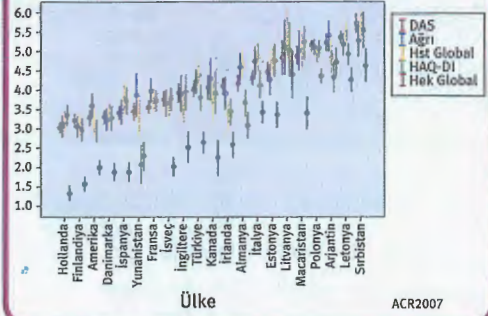
Günlük aktivitelerdeki yardım ihtiyacını da analiz ettik. Orijinal HAQ malüliyet skoru, "saçını taramak" veya "giyinmek" veya "yürümek" ve diğer günlük yaşam aktivitelerini ne kadar kolay yaptıkları ile ilgili 21 soruyu içerir. Orijinal HAQ aynı zamanda cihazların kullanımı ve her aktivite için diğer insanlardan istenen yardım ihtiyacını da sorar. Skorlamada, dahil edilen cihazlarla birlikte veya değil, HAQ skorlamasındaki farkı analiz ettik. Sonuçlar şöyle: 0.22'lik HAQ, klinik açıdan minimal kabul edilir. Yardım ve cihazlı veya yardım/cihazsız HAQ arasındaki fark sadece bir ülkede 0.22 ve diğerlerinde daha azdır. Benim gibi HAQ hesaplayanlar, "yardım/cihazları" dahil etmeden HAQ hesaplanması daha kolay olduğundan bundan çok memnun olacaklardır.

Yazarları arasında bulunduğum bir yayında, farklı ülkelerde kanserde 5 yıllık sağkalımla ilgili olarak ne yapıldığını görmek benim için ilginçti. Makro ekonomik değişken olarak gayrisafi milli hasıla kullanılmıştı. Çoğu ülkedeki kanser veri tabanları bu analizlere olanak sağlar ve muhtemelen diğer kronik hastalıklarda bu analizler yapılmıyordu.

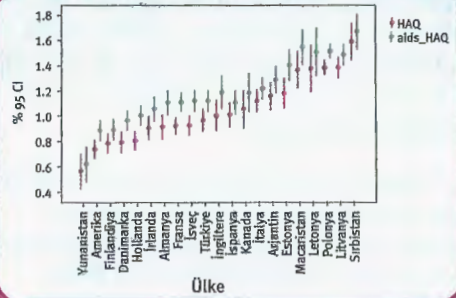
Daha yüksek GSMH olan ülkelerde 5 yıllık kanser sağkalımının daha iyi olduğu gösterilmiştir. Bu çok makuldür ve genelde yaşam beklentisi ve sağkalım dahil diğer verilere göre de GSMH ile bağlantılıdır.

QUEST-RA'da 20 ülkeden hastalık aktivitesi saptanmış ve GSMH ile ilişkisi grafikte gösterilmiştir. Burada yüksek GSMH olan ülkelerde ortalama DAS28 düşük ve düşük GSMH olan ülkelerde ortalama DAS28 yüksektir. Şekildeki yuvarlaklar, ülke başına ulusal sağlık harcamaları göstermektedir. Bu slaytla ilgili diğer bir konu ise şudur; Türkiye'deki hastalık aktivitesinin daha yüksek olması beklenirdi

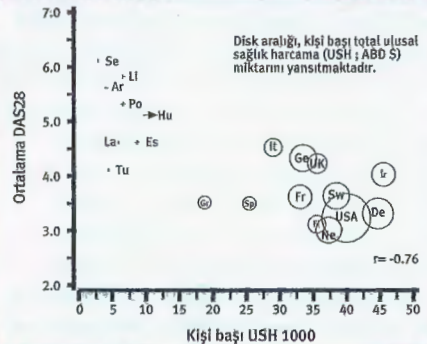
DAS 28, hasta tarafından bildirilen sonuçlar (PROs) ve QUEST-RA'da hekim global ölçümü



HAQ SKORLAMASI-NE KADAR GEÇERLİ
"diğer insanlardan, yönlendirmelerden
ve cihazlardan yardım"



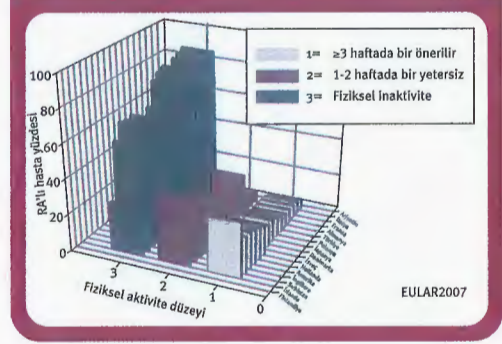
Makro ekonomik değişkenlerle
RA'nın klinik ilişkisi



ancak Türkiye’de beklenenden daha düşük hastalık aktivitesi bulunmuştur.

Fiziksel egzersize gelince; hasta anketinde en az yarım saat boyunca yüksek nabız hızında ve terleme vs. ile kaç kere aerobik, egzersiz yaptıklarını sorduk ve “haftada üç veya dört” ile “ayda bir ila iki” cevap seçeneklerimiz vardı. Haftada önerilen fiziksel egzersiz sayısı üç kez veya daha fazladır. RA’li hastalarda, yıllardan beri hastaların eklemlerini kullanmalarıyla ve egzersizin eklemleri ile ilgili sorunlar yaratabileceğinden duydukları endişeyle ilgileniyoruz. Onbeş ülkenin yer aldığı bu kesitsel çalışma, çoğu ülkede hastaların sadece % 10 ila % 20’sinin önerilen sıklık olan haftada üç veya daha fazla kez egzersiz yaptıklarını göstermektedir. Finlandiya’da RA’li hastalarımızın neredeyse üçte biri bu seviyede egzersiz yaptıklarını bildirir. Diğer çalışmalardan egzersizin sağlığa yararlı olduğunu bildiğimizden, hastalarımızın egzersiz alışkanlıklarını düşünmemiz gerektiğini düşünüyoruz.

QUEST-RA’da olduğu gibi, gözlemle daha iyi hasta bakımı yönünde yol alabileceğimizi ifade etmek isterim. Bu program aynı zamanda pek çok ülkedeki RA bakımından da ilginç bir özet sunuyor. Evrensel bir veri tabanı kurarak RA hastaları ile ilgili bilgi toplamak gibi bir hayalimiz var. İleride aynı zamanda klinik verilerle bir biobankanın oluşturulma olasılığı da vardır.





Konuşma Özeti*

“Kanıt dayalı tıp” kanıtları, sadece klinik çalışmaların sonuçlarıyla değil, klinik gözlemsel verilerin de kullanılmasıyla elde edilebilir. RA’te iyi bir veri tabanı oluşturmak için sadece anti-TNF ajan alan hastaları kaydetmek yeterli değildir. Kontrol amacıyla tüm hastaların verileri toplanmalıdır, aksi takdirde diğer hastaların ne gibi tedaviler aldığını; örneğin malignite gelişip gelişmediğini bilmek mümkün değildir.

Amerikan Ulusal Veri Tabanı, 1998’de Kansas’ta Dr. Wolfe tarafından araştırma ve eğitim amaçlı olarak oluşturulmuştur. Ardından anti-TNF tedaviler kullanılmaya başlanmış ve FDA pazarlama sonrası veri tabanı gözetimi olmasını zorunlu tutmuştur. Amerika’da İskandinav ülkelerinden farklı olarak kişisel gizliliğe çok önem verilmektedir. Dr. Wolfe’un bu konuda kullandığı ve oldukça başarılı olan yöntem, veri için primer olarak hastayla temas içinde olmaktır. Doktorun yapması gereken tek şey hastaya veri tabanına girmeyi isteyip istemediğini sormaktır. Kayıtların takibi aslında hastanın sorumluluğundadır. Eğer istemezse göndermeyebilir, ancak genellikle hastalar sisteme katkıda bulunmaktan mutludur ve ayrıca bu maliyeti en düşük seçenektir.

Dr. Pincus bunun yanında Nashville’de, 1982 yılında başlattığı ve 25 yıldır kullandığı kendi veri tabanlarından da söz etmiştir. Bu sistem, hasta anketlerinin röntgen filmlerinden ve laboratuvar tetkiklerinden çok daha fazla etkin olduğu gözlemi yapıldıktan sonra başlatılmıştır. Kliniğe geldiğinde hastaya iki sayfalık bir anket formu verilmekte ve her şeyin temelinde bu değerlendirilmektedir.

Klinik araştırmalarda kullanılan anketlerle bu sistemde kullanılan anketler arasındaki farkı Dr. Pincus şöyle vurgulamıştır: “bu sistemde kullanılan anketler 10 dakikadan kısa sürede doldurulabilecek anketlerdir. Hekime zaman kazandırır ve skorlama modelleri vardır. Vizitlerin yarısından fazlasında 28 eklem sayımı bile kullanılmamaktadır”.

Yirmisekiz eklem sayımını tamamlamak için, gerekli medyan süre 90 saniyedir; HAQ’ı tamamlamak için 41,9 saniye, MDHAQ RAPID3’ü tamamlamak için 10 saniye ve eğer kendi

ABD’DE ROMATİZMAL HASTALIKLAR İÇİN VERİ TABANLARI

Theodore Pincus

*New York Üniversitesi,
New York, ABD*

kendine bildirilen bir eklem sayısı ile birlikte yapılırsa 20 saniyedir. RAPID₃, üç ana veri setinden oluşur. Farklı metodlarla benzer veriyi elde ediyorsak çabuk sonuç vereni tercih etmek anlamlı olacaktır.

MDHAQ'ın ne olduğunu da açıklayan Dr. Pincus, MDHAQ'ın klasik HAQ'tan standart klinikte kullanılabilirliği artırmak için adapte edildiğini anlatmıştır. MDHAQ'ta aktivite sayısı 20'den 10'a indirilmiştir. MDHAQ fiziksel fonksiyon, ağrı, global durum, yorgunluk, hastanın bildirdiği eklem sayısı için her biri 0 ile 10 arasında olmak üzere skorlama modelleri ve ayrıca birleşik RAPID modelleri içerir ve klinik koşullarda 10-20 saniyede skorlanabilir. On cm'lik çizgi yerine 21 halkadan oluşan VAS kullanılmıştır ki bu, 5 saniye daha kazandırmaktadır.

Sayfanın bir yüzünde hastanın kendi kendine bildirdiği tüm ölçümlerle ilgili veri setleri ve skorlama modelleri, ikinci yüzünde ise sistemlere göre semptomların listesi ve yeni medikal öykünün gözden geçirilmesi yer alır.

Hasta anketleri iş görememezlik, maliyet ve prematür mortalite gibi uzun dönem RA sonuçları açısından en önemli klinik prognostik belirteçlerdir.

Konuşmasının sonunda Dr. Pincus, RA hastalık aktivite indekslerine de değinmiştir. Uzun dönemli gözlemler için indeksler kullanılmalıdır. ACR setini ve SDAI yani *simplified disease activity index* gibi ana veri setinden 5 ölçümü ve ağrıyı içermektedir ve DAS gibi bir formülü yoktur. Clinical Disease Activity Index, Viyana'da oluşturulmuştur ve ESR ve CRP'yi kaldırmıştır. Hatta Viyana'da bu ölçümler, klinik kararları vermede kullanılmamaktadır. RAPID skoru, hasta tarafından bildirilen üç ölçümü (fonksiyon, ağrı ve global tahmini) ölçer. RAPID₅ doktorun global tahminini de içerir. RAPID₂ ise sadece iki global veriden oluşur ve kıyaslamada en yüksek rölatif etkinliğe sahiptir. DAS kategorileri ile RAPID kategorilerini karşılaştırsak DAS skoru 5.1'in üstünde olan hastaların % 73'ünde RAPID skorunun 4'ten büyük olduğunu, yani oldukça yüksek olduğunu görmekteyiz, bu da RAPID skorları ile DAS skorlarının korele olduğunu göstermektedir.

Konuşma özeti Hasan Yazıcı tarafından ses kayıtlarından hazırlanmıştır.

HASTA İZLEMİ VE VERİ TOPLANMASININ GÜNLÜK PRATİĞE ALINMASI-ARMD VE BOARD VERİ TABANLARI

Yusuf Yazıcı

Günümüzde klinikte romatoloji hastası takibi hikaye, muayene ve bazı laboratuvar testlerinin yardımı ile yapılmaktadır. ABD’de romatologlara sorulduğu zaman % 15’ten daha azı her vizitte 28 eklem sayımı ve yine % 10’dan daha azı hasta formları kullanmaktadırlar. Bununla beraber yapılan laboratuvar araştırmalarından sedimantasyon ve C reaktif protein seviyeleri hastalık başladığında % 40’a varan oranlarda normal olabilir ve en etkili tedavi radyolojik hasar meydana gelmeden başladığında olmaktadır. Bu durumda hasta ve tedavi hakkında verilen birçok karar objektif olmayan ve/veya çoğu zaman hastalık aktif olmasına rağmen normal olabilecek verilere dayandırılmaktadır.

Herhangi bir kardiyoloji hastasının takibinde tansiyon ya da kolesterol veya tiroid hastasının TSH seviyesinin bakılmamış olabileceğinin düşünülmemesi gibi, ateşi olan bir hastanın sadece ateşi olduğunu bilmek yetmeyip tam derecesinin de ne olduğunu bilmek hastanın tedavisi açısından yararlı olacaktır. Romatoid artrit (RA) hastalarına gelince elimizde bunu sağlayabilecek, hastaların şu anda nasıl olduklarını objektif olarak gösterecek birçok yöntem olmasına rağmen bunlar klinik ortamında sıklıkla kullanılmamaktadır.

Birçok RA hastasının ağrılı eklemleri olduğu, kendini iyi hissetmediği hastayı dinlemek ve muayene ile anlaşılabilir. Ancak bunun ölçülüp, yazılması, zamanla ve tedavi ile ne kadar değiştiğinin gösterilmesi hem daha sağlıklı kararlar vermemiz hem de eğer bir gün bazı soruları cevaplamamız için çok yararlı bir veritabanı olacaktır.

ABD’deki romatologlara neden hasta formlarını kullanmadıkları sorulduğu zaman “çok zaman alıyor”, “hasta yapmak istemiyor” “hemşire /sekreter vermek istemiyor” “ölçümün yararlı olduğuna inanmıyorum” gibi bu gibi formları ya daha önce kullanmamış olmaktan doğan ya da sadece klinik ilaç çalışmaları sırasında kullanmış olmaya bağlı bir çekingenlik ortaya çıkmaktadır. Bunun yanında Avrupalı romatologlara da hangi sıklıkla 28 eklem sayımı yaptıkları sorulduğunda yarısından fazlası bunu yapmadıklarını belirtmişlerdir. Hem ABD hem de Avrupa’da RA hastalarını en iyi şekilde takip edebilmemiz için bazı değişikliklere gitmemiz gerektiğini düşünüyorum.

Rutin klinik ortamda kullanılan formlar hasta için doldurulması kolay, doktora zaman kazandıran ve aynı zamanda hasta takibi için yararlı bilgileri içeren formlardır. Bu hasta formları birçok ortamda denenmiş ve yararlı oldukları, objektif olarak hasta takibini sağladıkları, hatta laboratuvar testleri ve/veya eklem sayımını gerektiren hastalık aktivitesi indeksleri ile aynı derecede hastalık aktivite değişikliği gösterdikleri yerleşmiş ölçüm metotlarıdır.

Bizim hem özel klinik hem de hastane polikliniklerinde kullandığımız *multi dimensional health assessment questionnaire* (MDHAQ) 1980'lerin başında geliştirilmiş olan HAQ formunun rutin takipte daha kolay kullanılabilmesi için geliştirilmiş ve kısaltılmış halidir. Fonksiyon ile ilgili 10 sorunun yanı sıra ağrı, sabah sertliği, yorgunluk, hastanın *global değerlendirmesi*, kullanılan ilaçlar, *sistemlerin gözden geçirilmesi* ile ilgili başka hastanın takibinde yararlı olabilecek bilgiler de eklenmiştir.

Hatırlanması gereken; bu formların klinik araştırmalarda kullanılanlardan daha farklı amaçlar için geliştirildikleri ve bunun için daha kısa, normal hasta doktor ilişkisini etkilemeyecek ve asıl amaçlarının doktorun işini kolaylaştırmak olduğudur. Hastanın bu formu doldurması genelde 5 dakikadan az sürüp, doktorun kısaca göz atması 10 saniyeden kısa sürmektedir. Üzerindeki skorlama şemaları ile hesaplanan RAPID indeksi de doktor ve hastanın takip edebileceği bir hastalık aktivite skoru verebilir.

Pratikte başarı ile kullanılabilmesi için bazı yöntemlerin uygulanmasında fayda vardır. Ofis çalışanlarının, sekreterlerin hangi hastanın neyi olduğunu bilmeleri çok zordur. Aynı zamanda yeni gelen ya da semptomları ve/veya tanısı değişen bir hastada doktorun bile tam olarak ne olduğunu her zaman bilmesi mümkün değildir. Bununla beraber belli aralıklarla da formun verilmesi düşünülürse -mesela her 6 ayda bir, ya da senede bir- bu sefer de hangi hastanın ne zaman en son formu doldurduğunu takip etmek ve ona göre davranmak zor olacaktır. Pratikte işleyen tek yöntem her hastanın her vizitte bu formu doldurmasıdır. MDHAQ ile toplanan bilgilerin bir çok romatolojik hastanın takibinde yararlı olduğu gösterilmiştir.

Ofiste çalışanlara bu formların ne kadar önemli olduğu anlatılmalıdır. Eğer formu hastaya verecek kişi bunun doktor için önemli olmadığını, sadece yapmış olmak için yapıldığını düşünürse kısa sürede formun hastalara verilmesi aksayacak ve sistem oturmayacaktır.

Görülen her hastadan her seferinde bu verilerin toplanması aynı zamanda araştırma yapmak isteyen doktorlar için de bulunmaz bir fırsattır. Her hasta dahil edildiği için bazı veritabanlarında görülen "seçilmiş" hasta sorunu bir derece azaltılmış ve aynı zamanda hastalıklı kontrol grupları da kendiliğinden toplanmış olur.

Başka önemli bir nokta formun olabildiğince hasta tarafından yardım almadan doldurulmasıdır. Hemşire, sekreter yardımı olduğu zaman ikinci bir gözlemci ve onun yorumu cevaplara yansımaktadır, en iyi yöntem hastanın en doğru cevabın ne olduğunu düşünüyorsa onu yazmasıdır.

MDHAQ formu birçok kısımdan oluşur ve bunlar aşağıda gösterilmiştir (Şekil 1). Bunlar arasında fonksiyon, ağrı ve hastanın global değerlendirme skorları kullanılarak oluşturulan RAPID indeksi (Şekil 2) DAS28 ile korelasyon gösterip pratikte kullanılması çok daha kolay olan bir indekstir. RAPID'i oluşturan bu 3 kısmın özellikle doldurulmuş olmasına dikkat etmek gerekir.

RAPID 0-30 arası skorlanan bir indekstir. Fonksiyon 0-10, ağrı 0-10 ve hastanın global değerlendirmesi skorları 0-10 skorlanıp üçü toplanarak RAPID skoruna ulaşılır. Üç ve altındaki skorlar remisyon, 3-6 arası düşük aktivite, 6-12 orta ve 12 ve üstü skorlar yüksek seviyede aktivite şeklinde yorumlanıp ona göre tedavide değişiklik ya da aynı tedaviye devam düşünülebilir. Unutulmaması gereken her indeksin hastalık ile direkt ilgisi olmayan başka olgulardan etkilenebileceğidir. Örneğin bir travma sonrası sırt ağrısı olan hastanın ağrı skoru RA'e bağlı olmadan yüksek olup RAPID skoru da buna bağlı olarak yüksek olacaktır. Ancak bütün bunların hastanın muayene ve anamnezden toplanan bilgiler ile bir bütün olarak yorumlanması ve tedavi hakkında kararlar verilmesi gerekir.

Hasta formlarının uzun süreli sonuçlar ile ilişkisine bakıldığı zaman laboratuvar ve radyolojik verilerden daha fazla mortalite, eklem cerrahisi, işe devam, fonksiyon ve maliyet ile ilişkili ve prognostik olarak daha güçlü oldukları görülmüştür.

Pratikte veri toplamanın en önemli nedeni ise kanımca bu tip skorlar ile takip edilen ve tedavileri bu objektif skorlara göre değiştirilen hastaların bu gibi skorlar ile takip edilmeyen hastalara göre daha iyi oldukları ve fonksiyonlarını daha iyi koruduklarının görülmüştür.

Buna en iyi örnek 1990'ların sonunda yapılan TICORA çalışmasıdır. Hastaların yarısı rutin takip yarısı da her ay hesaplanan DAS28 skoruna göre tedavi edilmiş ve 1 sene sonunda DAS28 ile takip edilen hastaların % 65 remisyonda olduğu (% 16 kontrol rutin takip grubunda) görülmüştür. Bu hastaların sulfasalazin ile tedavi edildiği ve

son 7-8 senede çıkan biyolojik ajanlar ile tedavi edilmedikleri de göz önüne alınırsa objektif skorlar ile yakın takibin hasta sonuçlarında kullanılan ilaçtan daha önemli, en azından ilaçlar kadar önemli olduğu sonucu çıkarılabilir. Daha sonra tamamlanan BeSt çalışması da buna benzer sonuçlar vermiştir.

MDHAQ ile toplanan verilerin hasta dosyasında kalmasının yanı sıra bir veri tabanına girilmesi de düşünülebilir. Asıl amacı klinik çalışmalarda kullanılmak olan bu veri tabanı, günlük hasta takibinde de eski ve yeni bulguları kolayca gösterebildiği için yararlı olabilir. RA hastalarını daha objektif metotlarla takip ve tedavi etmemiz gerektiği bu şekilde tedavi edilen hastaların daha iyi sonuçlar vermeleri ile gösterilmiştir. MDHAQ bunun genel pratikte kolayca uygulanabilecek bir yoldur.

1. Lütfen şu anda kabiliyetiniz ile ilgili en doğru BİR cevabı (✓) ile işaretleyiniz:

SON BİR HAFTA İÇERİSİNDE, aşağıdakileri yapabildiniz mi:	HERHANGİ BİR zorluk olmaksızın	BİRAZ zorlukla	BÜYÜK zorlukla	YAPAMADIM
a. Giyinebildiniz mi? (ayakkabılarınızı bağlamak ve düğme ilikleme dahil)	☐ 0	☐ 1	☐ 2	☐ 3
b. Yatağa girebildiniz mi ve yataktan çıkabildiniz mi?	☐ 0	☐ 1	☐ 2	☐ 3
c. Dolu bir fincan ya da bardağı ağızınıza götürebildiniz mi?	☐ 0	☐ 1	☐ 2	☐ 3
d. Dışarıda düz zemin üzerinde dolaşabildiniz mi?	☐ 0	☐ 1	☐ 2	☐ 3
e. Tüm vücudunuzu yıkayıp kurulaşabildiniz mi?	☐ 0	☐ 1	☐ 2	☐ 3
f. Yerden giysilerinizi kaldırmak için eğilebildiniz mi?	☐ 0	☐ 1	☐ 2	☐ 3
g. Normal muslukları açıp kapatabildiniz mi?	☐ 0	☐ 1	☐ 2	☐ 3
h. Bir araba, otobüs, tren ya da uçağa binip inebildiniz mi?	☐ 0	☐ 1	☐ 2	☐ 3
i. İsteddiğinizde iki mil ya da üç kilometre yürüyebildiniz mi?	☐ 0	☐ 1	☐ 2	☐ 3
j. İsteddiğinizde, eğlenceli aktivitelere ve spora arzu ettiğiniz kadar katılabildiniz mi?	☐ 0	☐ 1	☐ 2	☐ 3

FN

☐	1=0.3	5=1.7	9=3.0	13=4.3	17=5.7	21=7.0	25=8.3
	2=0.7	6=2.0	10=3.3	14=4.7	18=6.0	22=7.3	26=8.7
	3=1.0	7=2.3	11=3.7	15=5.0	19=6.3	23=7.7	27=9.0
	4=1.3	8=2.7	12=4.0	16=5.3	20=6.7	24=8.0	28=9.3
							30=10

2. Durumunuz nedeni ile SON BİR HAFTA İÇERİSİNDE ne kadar ağrınız oldu?
Lütfen aşağıda ağrınızın ne kadar şiddetli olduğunu belirtin:

AĞRI ○ OLABİLDİĞİNCE
YOK 0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10 KÖTÜ BİR AĞRI

PN

3. Hastalığınız ve sağlık koşullarının sizi şu anda etkileyebileceği tüm durumları düşünerek,
lütfen aşağıda nasıl olduğunuzu belirtin:

ÇOK ○ ÇOK
İYİ 0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10 KÖTÜ

RAPID

(0-30)

DAS28, RAPID ve ACR RA
cevap indekslerini oluşturan kısımlar

	DAS28	RAPID	ACR
Şiş eklem	+		+
Hassas eklem	+		+
Doktor global skoru			+
ESR/CRP	+		+
Hasta global skoru	+	+	+
Fonksiyon skoru		+	+
Ağrı		+	+

REFERANSLAR: [1, 2]

1. Pincus T, Yazıcı Y, Bergman M: Development of a multi-dimensional health assessment questionnaire (MDHAQ) for the infrastructure of standard clinical care. *Clinical and experimental rheumatology* 2005, 23 (5 Suppl 39):S19-28.
2. Yazıcı Y: Databases in routine care: possible and necessary. *Bulletin of the NYU hospital for joint diseases* 2007, 65 (2):127-131.
3. Pincus T, Segurado OG: Most visits of most patients with rheumatoid arthritis to most rheumatologists do not include a formal quantitative joint count. *Annals of the rheumatic diseases* 2006, 65 (6):820-822.
4. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, Kincaid W, Porter D: Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004, 364 (9430):263-269.
5. Pincus T, Yazıcı Y, Bergman M: Saving time and improving care with a multidimensional health assessment questionnaire: 10 practical considerations. *The Journal of rheumatology* 2006, 33 (3):448-454.
6. Pincus T, Callahan LF: Rheumatology function tests: grip strength, walking time, button test and questionnaires document and predict longterm morbidity and mortality in rheumatoid arthritis. *The Journal of rheumatology* 1992, 19 (7):1051-1057.
7. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJ, Hazes JM, Zwinderman AH, Peeters AJ, de Jonge-Bok JM, Mallee C *et al*: Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. *Annals of internal medicine* 2007, 146 (6):406-415.
8. Yazıcı Y: Monitoring response to treatment in rheumatoid arthritis--which tool is best suited for routine "realworld" care? *Bulletin of the NYU hospital for joint diseases* 2007, 65 Suppl 1:S25-28.

[3]

[4-8]

Herkese Günaydın,

Bugünkü interaktif oturum istatistik güç analizi konusunda bir uygulama olacak. Dün konuştuğumuz konuların günlük hayatımıza kolay girebilmesi amacıyla, internetten ücretsiz indirebileceğiniz bazı programları kullanarak güç analizi alıştırmaları yapacağız. Aramızda genetik ilişki çalışmaları yapanlar da olduğundan, klinik ilaç çalışmaları için yapacağımız alıştırmalardan sonra, genetik araştırmalarda güç analizi uygulamaları da olacak. Ama uygulama öncesinde Schering Plough Türkiye'ye teşekkür ederek başlamak istiyorum. Hem geçen seneki toplantı, hem onun devamı olan bu toplantıda, klinik pratiğimize nesnel tıbbın yerleşmesi için temel teşkil edebilecek ana konuları özgürce işleyebileceğimiz bir ortamı bizlere sundular. İlk toplantının Bodrum'da yapılması belki bir tesadüftü. Fakat ikincisinin de Bodrum'da olması bunun artık gelenek olabileceği yolunda kuvvetli şüpheler oluşturuyor. Umarım Bodrum toplantıları gelenekselleşir ve nesnel tıbbın günlük pratiğimize girmesi açısından oldukça yararlı bir açılım olur.

Bu uygulamanın rahat bir ortamda olmasını istiyoruz. Dört gruba ayrılarak ön sıralarda yer alan 4 bilgisayar ile çalışacağız. Bilgisayarlarda 3 web sayfasının açık olduğunu göreceksiniz. Ayrıca, 2 tane program da desktop üzerinde duruyor. Uygulamada bu programları kullanacağız ve elimizdeki örnekleri kullanarak simülasyon denemeleri yapacağız. Daha sonra elde ettiğimiz sonuçları birlikte tartışacağız. Bu bir tür *dry lab* uygulaması olacak. Buradaki amacımız dün konuşulan konuların uygulamasını yapmak. Ben sadece bir aracı olacağım. Amacımızı bir Nasrettin Hoca fıkrası üzerinden anlatmak da mümkün. Hocaya soruyorlar, hocam un var mı? –Var; Şeker var mı? –Var; Yağ var mı? –Var. Ya Hocam o zaman niye helva yapmıyorsun? Şimdi bizim kafamızda genellikle helva yapmak var. Amacımız helva yapıp bu helvayı mümkün olan en güzel yerde sunmak. Fakat genellikle helva yapmak istediğimizde evimizde ya yağ olmuyor, ya irmik olmuyor, ya da unumuz az oluyor. Ama biz her durumda bir helva yapıyoruz. Bunun adı helva da olsa tadı farklı olabiliyor ve genellikle de istendiği gibi iyi sunulmuyor. Bu toplantının amacı eğer gerçekten helva, “iyi bir çalışma” yapmak

İTERAKTİF OTURUM I

Ahmet Gül

*İstanbul Üniversitesi,
İstanbul Tıp Fakültesi, İstanbul*

istiyorsak, bunun bütün aşamalarını baştan planlamamız gerekiyor. elimizde helva ile ilgili reçete neyse, o reçeteye uyacak şekilde malzemeleri toplamamız gerekiyor. Günümüzde çalışmaların yeterli istatistik güce sahip olmasını ve yeterli sayıda örneği ve malzemeyi bulmak istiyorsak, sıklıkla işbirliği yapmamız gerekiyor. Bunu özellikle klinik ilaç araştırmalarında görüyoruz. Klinik ilaç araştırmalarının büyük kısmının çok merkezli katılım olmadan yapılamayacağını biliyoruz. Bu işbirliğinin aslında en önemli nedeni, yeterli istatistik güce ulaşabilmek için uygun sayıda örneğin çalışmaya dahil olabildiğini sağlamak. Bunu ilaç çalışması dışındaki araştırmalarda da görebiliyoruz. Sizin elinizde yağ yoksa komşudan alacaksınız, elinizdeki unun miktarı az ise, bunu gidip bir başkasından alarak tamamlayacaksınız. Eğer amacınız gerçekten helva yapmaksa, bütün içeriği yeterli miktarda olmalı. Etik açıdan düşünüldüğünde, bir çalışmada hipotezinizde öngördüğünüz sonuca ulaşacak gücü sağlayandan daha az sayıda örneği çalışmak da hatalıdır, gereğinden fazla örneği çalışmak da. İşin artan maliyeti ötesinde, etik boyutunu da düşünmek gerekir.

Klinik ilaç araştırmaları için endüstri desteğini, bu araştırmaların toplam bütçelerini ve güç analizinin önemini göz önünde alınca, bu amaçla kullanılabilecek ve ciddi ücretler karşılığında satın alınabilecek bilgisayar programlarını görmek şaşırtıcı değil. Fakat internette kısa bir araştırma yaptığınızda, gerçekten işe yarayan ve ücretsiz olarak kullanılabilen çok güzel programları bulmak da mümkün. Bunlardan bir kaç örneği birlikte kullanacağız. Ardından genetik ilişki çalışmaları ile ilgili uygulama yapacağız. Doğrusu, bu uygulamalar sırasında, programları kullanmak kadar, bir çalışmayı planlarken hangi özellikleri göz önüne almamız gerektiğini görmemizin bile büyük yarar sağlayacağını düşünüyorum.

Şimdi bilgisayarlarımızı açarsak, uygulama dosyasının içerisinde programları bulacağız. Bunlardan bir tanesi Clinic Trial Simulator (CTS). Bu program randomize kontrollü klinik çalışmalar için simülasyon yoluyla güç analizi yapmayı sağlayan ve ücretsiz olarak bilgisayarınıza indirebileceğiniz bir yazılımın beta versiyonu (www.randomization.org).

Bir AB projesi olan PRACTIHC (*Pragmatic Randomized Controlled Trials in HealthCare*) tarafından desteklenmiş (www.practihc.org). Program bize iki konuda yardım sağlıyor: 1- Gerekli örnek grubu büyüklüğünü hesaplıyor, çalışmayı ideal koşullarda nasıl yapmamız gerektiğini gösteriyor. 2-Ardından klinik araştırmanın binlerce kez (örneğin 5000 kez) simülasyonu yapabiliyor. Siz istediğiniz parametreyi değiştirerek ideal örnek grubunda beklediğiniz sonuçlarla simüle edilmiş örneklerdeki farkı test edebiliyorsunuz.

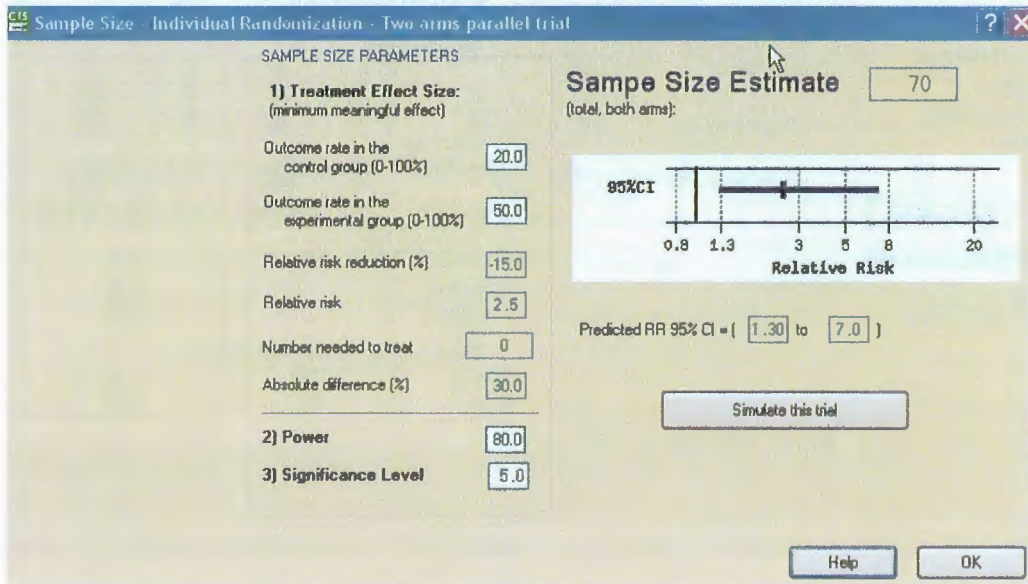
Programda örnek büyüklüğünü hesapla (*compute sample size*) bölümüne girelim. Burada *individual randomization* bölümüne giriyoruz. İki kollu bir paralel çalışmada, standart formül kullanılarak gerekli örnek grubu büyüklüğünü size veriyor. Burada çıkan rakamlar tümüyle ideal olan bir çalışmanın size verebileceği sonuçlar. Burada elde ettiğiniz verileri daha sonra simüle etme şansımız var. Şimdi örnek büyüklüğü parametrelerine bakalım. Plasebo kontrol grubunda beklediğimiz yanıt oranı için bir örnek seçelim. Behçet hastalığında oral aft tedavisi çalışması yaptığımızı varsayalım. Plasebo kullandığımızda oral afta düzelme olarak beklediğimiz yanıt % 20 olsun. Deney grubuna da ilaç veriyoruz ve bu ilaçta da beklediğimiz yanıt % 50 olsun. Dikkat ederseniz plasebodan % 30 farklı yanıt varsayıyorum. Aşağıda, dün konuştuğumuz istatistik güç beklentileri var. Burada gücü % 80 olarak hedeflemişiz ki genellikle kabul edilebilir bir sınır olarak görülür. Beta değeri olarak % 20'yi kabul etsek, 1beta yani % 80 güç ile çalışmayı yapacağız demektir. Ardından istatistik anlamlık derecesini girmemiz gerekiyor. Bunu % 5 olarak belirledik, ama bunu % 1 üzerinden de hesaplayabiliriz. Bir de beklediğiniz rölatif risk (RR) skalası (*relative risk scale*) var. Hangi RR skalasında sonuç alabileceğinizi gösteriyor (0 ila 2 arası gibi). Çünkü konuşmaları hatırlarsanız, eşitlik çalışmasında ya da noninferiority çalışmasında beklenen RR yerinin de kayması söz konusu olabilir. Verileri girdiğimde karşıma çıkan rakam 70. Eğer çalışmaya 70 hasta alırsam ve 2 kola eşit olarak dağıtırsam, ideal koşullarda benim % 80'e yakın bir güçle anlamlık sınırına ulaşan bir veri elde etme şansım var. Ama tamamen ideal koşullardan bahsediyoruz: 35 hasta bir kolda, 35 hasta diğer kolda, % 20 yanıt

plasebo kolunda ve % 50 yanıt ilaç kolunda, ayrıca iki kol da birbirine her şeyiyle eşit koşullarda, çalışmayı terkeden kimse yok. Fakat gerçek dünyada koşullar hiçbir zaman böyle olmuyor. İsterseniz bu aşamada birkaç egzersiz yapalım. Yani farklı koşulları deneyelim ve daha sonra da işin ikinci kısmına geçelim.

Simulasyon bölümünde gerçek dünyada gördüklerimize benzer sorunların etkilerini test edebilme şansımız olacak. Plasebo ve ilaç kollarının eşit olup olmaması, hastaların tedavi uyumu ya da takipten çıkmalar gibi sorunların güç analizine etkilerini göreceğiz. “*Simulate this trial*” tuşuna bastığımızda önümüze belirli sayıda simule edilmiş bir sonuç karşımıza çıkacak. Sol sütunda “*True effect size and N*” sekmesini göreceğiz. Bunun altında “*true effect size*” bölümüne girerek beklediğimiz etki büyüklüğünü değiştirme şansımız olur. Bu sayede istediğimiz türde bir ilaç çalışmasından (üstünlük ya da eşitlik gibi) beklediğimiz etki büyüklüğünü girme sonucu görme şansımız var.

“*Number of subjects*” bölümünden de hastaların kollar arasındaki dağılımını değiştirebiliyoruz. Alttaki “*Trial problems*” sekmesi altında da takipten düşen (*Lost to Followup*) ve uyumsuz (*NonCompliance*) hastalarla ilgili değişiklikleri girmek mümkün. Bu değişiklikler sonrasında tekrar “*Run simulation*” tuşuna basınca karşımıza yeni sonuçlar çıkmakta. Yeni koşulların hangi oranda güç kaybına neden olabileceğini burada görebilmekteyiz. Dilerseniz bu sonuçları şekil üzerinde de görmek mümkün. Size rölatif riskin ve beklenen P değerlerinin simule edilen çalışma sayısı arttıkça gösterdiği değişiklikleri “*Figures*” sekmesi altında inceleme şansımız var. Bu sayede anlamlı sonuçlara ulaşma olasılığımızı da görsel olarak değerlendirebiliyoruz.

Eğer çalışmanızda altgrup analizi yapmayı hedefliyorsanız, güç analizini bu koşullara göre değiştirmeniz ve yapmak istediğiniz grup özelliklerine göre simulasyonu tekrarlamanız da mümkün. Benzer şekilde program farklı ilaç



araştırma koşulları arasında karşılaştırma yapma şansını da size tanıyor.

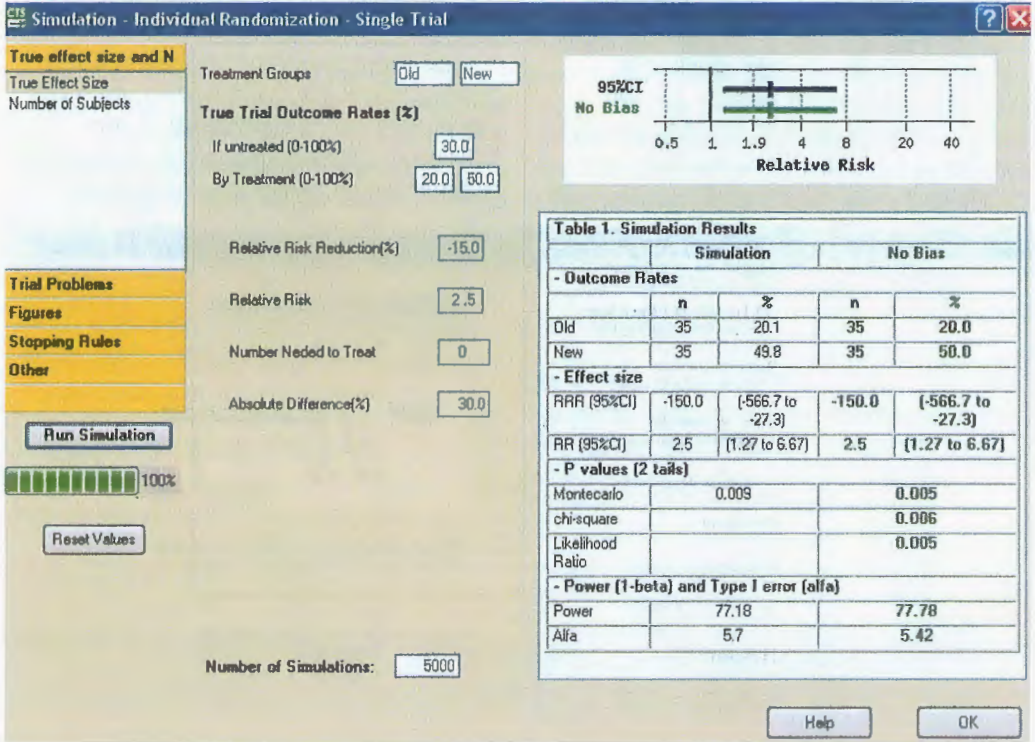
Soru: Aynı verileri de kullansak, her “run simulation” dediğimizde güç ve anlamlılık değerleri değişiyor.

Yanıt: Doğru, çünkü, her seferinde eldeki verilerle bir simülasyon yapıyor program. Bu nedenle 5000 kez yaptığı simülasyon sonucunda birbirine çok yakın, ama birebir aynı olmayan sonuçları almak mümkün oluyor.

Katkı: Ahmet’in gösterdiği çok güzel bir program. Şimdiye kadar kullandığımız güç hesabı yapan başka programlar da var. Ama

onlarda mesele sadece “örneklem büyüklüğü hesabı” olarak görülüyor. Onları kullandığınızda, sen kendi kafandan diğer hesapları yapıyorsun. Diyorsun ki, şu kadar kişi kaybolursa, şu kadar kişi almasa, filan 40 kişilik çalışma yapmak istiyorum ben, buna kaybedeceklerimi düşünerek 50 tane alayım. Bilgisayarın yaptıklarını sen kağıt kalemle yaptıktan sonra girdiğinde, sana sadece güç hesabını verenler var. Bu programın güzelliği senin kağıt kalemle daha önceden yapacaklarını da çok iyi yapıyor. O güzel bir şey.

Bu program gerçek dünyayı çok iyi yansıtıyor. Bakın bu sütunların sağda olanı, hep bizim programları kullandığımız zaman karşımıza çıkan rakam. Oysa çok sayıda örneği



simüle ettiğimizde çıkanlar, gerçek hayattaki beklentilerimize daha da örtüşüyor. Yani simüle zaman bu rakamların daha farklı olmasını bekliyoruz. Ve bunun kabaca dağılımını yansıtıyor. Yani aynı çalışma şu kadar kez yapılsa ne kadar farklı sonuçlarla karşılaşma olasılığımız olduğunu bize simulasyon ve şekillerle gösteriyor.

Kullanabileceğimiz bir diğer program PS: Power and Sample Size Calculation (version 2.1.31, 2004). Vanderbilt Üniversitesi tarafından geliştirilen bu program da ücretsiz indirilebiliyor:

<http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize>. Bu programda uygulamak istediğimiz istatistik testine göre veri girmek mümkün. Örneğin t testi için ayrı, χ^2 yöntemi kullanacaksanız (*dichotomous*) ayrı sekmeleri kullanmak gerekiyor. Aynı programda güç analizi ya da örnek büyüklüğü hesaplamak mümkün. Girmeniz gereken veriler benzer. Alfa değeri, hedeflenen güç, sağlıklı kontrollerde gözlenmesi beklenen olay ya da yanıt oranı (P_0), kontrol ve hastaların oranı (m), beklenen risk oranı (*odds ratio*, ψ) gibi. Ardından "calculate" tuşuna basınca sonucu hemen görebiliyorsunuz. Bu oldukça basit ve kullanışlı bir program.

Power and Sample Size Program: Main Window

File Log Help

Survival t-test Regression 1 Regression 2 **Dichotomous** Log

Output

Studies that are analysed by chi-square or Fisher's exact test

What do you want to know? Sample size

Design

Matched or Independent? Independent

Case control? Retrospective

How is the alternative hypothesis expressed? Odds ratio

Uncorrected chi-square or Fisher's exact test? Uncorrected chi-square test

Input

α p_0

power m ψ

Calculate

Graphs

Logging is enabled.

Exit

Önümüzdeki uygulama içinde başka bir örnek var: GPower. Bu bir Alman programı (<http://www.psych.uniduesseldorf.de/aap/projects/gpower/>) ve MSDOS temelli bir program. (Yeni hazırlanan GPower v3 Windows XP ve Vista sistemlerinde çalışıyor). GPower programında da farklı istatistik testleri kullandığınızda beklediğiniz gücü hesaplayabilme şansınız var. Yani t testi ya da χ^2 kullanmanıza göre, tek tek size beklediğiniz etki büyüklüğünü, alfa değerini, hedefinizdeki gücü ve istatistik anlamlılık sınırını sorup, buna göre bir hesaplama yapabiliyorsunuz.

Bu uygulamadan akılda kalması gereken en önemli nokta, klinik çalışma planlanırken hangi parametrelerin sonucu etkileyebileceğini bilmemiz ve güç hesabının da bu değişkenlerden ne kadar etkilenebileceğini görmemiz. Bu değerlendirmeler sonunda elde ettiğim verilere göre çalışma planladığımda, az da olsa başarısız kalma olasılığım var. Ama o zaman, başta yaptığım güç hesabımda bazı değişkenleri hesaba katmadığımı görebilirim. Veya bunun tam tersi şanslı çalışmalar da oluyor. Güç analizi yapmadığım halde, çok az sayıda örnekte o gücü yakalayabilme şansım olabiliyor. Ama bu çalışmada bana istatistik olarak anlamlı görünen bir sonuç vermekle beraber, buna benzer başka çalışmalar yapsam, çalışmanın gücü yeterli değilse anlamlı sonuç elde etme şansım çok düşük olacak. Buna benzer, başkaları tarafından tekrarlanıp doğrulanamayan pek çok çalışmadan bahsetmek mümkün.

Genetik İlişki Çalışmaları

Genetik araştırmalarda çok sık olarak vaka kontrol ilişki çalışması yapıyoruz. Hatırlarsınız bundan önceki toplantılarda özellikle ülkemiz için sağlıklı kontrollerin seçimi ile ilgili sıkıntılardan bahsetmiştim. Vaka kontrol çalışmalarında, kontrol seçimi çok önem taşıyor. Bununla birlikte, ideal koşullarda vakaları ve kontrolleri seçsek bile, eğer yeterli güç hesabını yapmamışsak, yaptığımız genetik çalışmanın sonuçları da iyi bir “helva”ya dönüşmüyor. Başlangıçta söylediğimiz gibi, iyi bir çalışma için önce elimizde yeterli malzemenin olması gerekiyor. Bunu nasıl yapabiliriz? Bu amaçla genetik ilişki çalışmaları öncesinde nasıl güç analizi yapabiliriz konusunu, Hasan Yazıcı

Hoca'nın da hoşuna gideceğini umarak, “genotipleme” hatasını da dikkate alarak gündeme getireceğim. Genotipleme hatasını Hasan Yazıcı hocamızla birlikte daha önce çeşitli vesilelerle tartışmıştık. Laboratuvar koşullarında hangi oranda genotipleme hatası yapıyoruz ve hatalı genotipleme sonuçlarımızı ne oranda etkiliyor bilmemiz gerekiyor. Özellikle günümüzde giderek artan oranda çok yüksek verimli (*highthroughput*) genotipleme platformları kullanılmaya başlandı. Bu sistemlerde bahsettiğimiz genotipleme rakamları yüzbinler, milyonlar düzeyinde. Dolayısıyla başlangıçta belirli bir hata oranını kabul etmek zorundasınız. Ön çalışmalarla elde edeceğiniz sonuçlara dayanarak, ne oranda hatalı genotiplemeyi göze alarak anlamlı bir sonuç hedefleyebilirim sorusunu cevaplamaya çalışacağız. Genetik ile ilgilenen arkadaşlar için <http://linkage.rockefeller.edu> web sayfasını kuvvetle tavsiye ederim. Bu Prof. Jürg Ott tarafından yönetilen ve Rockefeller Üniversitesi tarafından ev sahipliği yapılan bir genetik analiz sitesi. Jürg Ott en önemli genetik istatistikçilerden. Bu işle uğraşan pek çok insan onun talebesi, genetik bağlantı analizinde kullanılan pek çok program da, onun orijinal programından türetilmiş durumda. Bu sayfada genetik ilişki çalışmaları ve bağlantı analizi ile ilgili her türlü duyuruyu da bulmak mümkün. Ama en önemlisi tüm genetik analiz programlarına buradan ulaşmanın mümkün olması. Sayfada “Genetic Analysis Programs” altında, “A list of genetic analysis programs” yazısına tıkladığınızda karşınıza sık olarak güncellenen bir genetik istatistik programları listesi çıkıyor. Bugüne kadar genetik inceleme için kullanılan ne kadar program varsa buradan indirebilme şansınız var. Listede söz konusu programlar kimler tarafından yazılmış, özellikleri nelerdir, bağlantılı web sayfası neresidir, nasıl indirilebilir görebiliyorsunuz. Burada hangi genetik analiz programını uygulamak istiyorsanız hangi kaynak dilin kullanıldığına dikkat etmelisiniz. Çoğunlukla bu programların Unix ya da Linux için yazıldığını göreceksiniz. Eğer bu sistemle çalışan bilgisayarınız yoksa Windows ya da MSDOS tabanlı programlar bulmaya çalışmak gerekiyor.

Bu web sayfasının son kısmında PAWE (Power for Association with Error) diye bir program bulacaksınız. Derek Gordon tarafından

yazılmış olan bu programın özelliği, güç analizi yaparken ne kadar genotipleme hatasını göze alabilirsiniz sorusunu yanıtlamaya çalışması. Bu programı internet arayüzü ile kullanmak mümkün: “*Power for association with errors web interface*”. Bu sayfaya girdikten sonra “Run PAWE” satırını tıklayalım. Programın ihtiyacı olan verileri adım adım sorulan soruları yanıtlayarak girmemiz gerekiyor. İlk bölümde cevaplamamız gereken elimizde sabit sayıda örnek olduğunda ne kadar istatistik gücümüz olduğunu mu, yoksa arzu ettiğimiz gücü sağlayacak örnek sayısını mı öğrenmek istiyoruz sorusu olacak. Dilerseniz örnek sayısı hesabını seçelim.

İkinci olarak incelemeyi hangi modelle yapacağımızı soruyor. Eğer araştırdığınız hastalığın bir genetik geçiş modeli varsa, bunu hesaplamaya katmak mümkün. Uğraştığımız hastalıkların büyük çoğunluğu kompleks

hastalık olarak tanımlandığından, burada genetik modelden bağımsız, yani “*model-free*” hesaplamayı seçmeliyiz. Yani, planladığımız genetik ilişki çalışması için herhangi bir genetik geçiş modeli olmadığını varsayalım. Üçüncü soru genotipleme hatasını hangi yöntemle düzelteceğimizi soruyor. Bu benî de aşan bir soru, dolayısı ile yardım (*help*) bölümüne giriyorum. Farklı kişiler tarafından geliştirilen 4 tane genotipleme hatasını düzeltme yöntemi verilmiş. Birinin diğerine üstünlüğünü tartışacak durumda değilim. Fakat bunların herhangi birisini kullandığınızda, kabaca genotipleme hatasının getirebileceği sorunları yansıtır, size bir düzeltme tavsiye ediyor. Biz Derek Gordon ve Jürg Ott’un içinde olduğu Gordon Heath Liu Ott (GHLO) modelini seçelim. Bu sayfadaki son soru hedeflediğimiz anlamlılık düzeyi ile ilgili, Yani bize alfa değerini soruyor. Bunu da isterseniz 0.05 kabul edelim.

Specify the analysis you would like to perform ☐ Power calculations for a fixed sample size ☒ Sample size calculations for a fixed power	Help
Specify how to generate the genotype frequencies ☒ Genetic model free method ☐ Genetic model based method	Help
Specify the error model you would like to consider ☒ Gordon Heath Liu Ott ☐ Douglas Skol Boehnke ☐ Sobel Papp Lange ☐ Mote Anderson	Help
Specify the Significance Level of the test between 0 and 1 0.05	Help

Bir sonraki sayfaya geçtiğimizde ne kadar güç hedeflediğimiz soruluyor. Bu soruyu da isterseniz 0.80 olarak yanıtlayalım. Ardından çalışmamızdaki sağlıklı kontrollerle hastalarınızın oranını girmeliyiz. Kontrol ve hastaları eşit oranda alabilirsiniz. Hasta sayısını artırma şansınız yoksa, kontrol sayısını artırarak gücü artırabilmek mümkün olabilir. Önemli konulardan birisi de örneklerdeki alel dağılımının Hardy Weinberg dengesine (HWE) uyup uymadığı. Kullandığımız polimorfizmin sağlıklı kontrollerde HWE uyumunun olmasını, yani alel dağılımın uzunca bir zaman diliminde dengeye ulaşmış olmasını bekleriz. Sağlıklı kontrollerde HWE saptanmayan bir polimorfizmi incelemeye dahil etmemek gerekir. Ama hastalarda alel dağılımında HWE sapması beklenebilir. Bu nedenle sadece kontrollerde HWE beklediğimizi işaretleyelim.

Soru: Hastalarda her durumda HWE bozulmuş olması gerekmez mi? Yani bizim bir ilişki tanımlamamız için HWE sapması görmemiz

gerekmiyor mu?

Cevap: Bu söylediğiniz doğru olmakla beraber, ilişkinin büyüklüğüne bağlı olarak hastalıkla ilişkili alelin zaman içerisinde yeniden bir denge durumuna ulaşması da beklenebilir. Yani bu durum alelin dağılımına ve ne kadar zamandır hastalıkla ilişkili olduğuna göre değişkenlik gösterebilir.

Ardından ilk sayfada seçtiğimiz GHLO hata düzeltme modeline ait ϵ_1 ve ϵ_2 değerlerini girmemiz gerekiyor. Burada ϵ_1 değeri 1 alelinin hatalı olarak 2 aleli olarak belirlenme olasılığı, ϵ_2 değeri de 2 alelinin hatalı olarak 1 olarak belirlenme olasılığıdır. Bunları da her ikisi için 0.01 olarak girelim. Yine unutmamak gerekiyor, burada bir tanı hizmetinden bahsetmiyoruz, araştırma amacıyla yaptığınız genotiplemede hata beklentisinden bahsediyorsunuz. Tanı hizmeti veriyorsanız, % 1 de büyük bir rakam, ama biz burada bilimsel araştırmada karşılaşılabileceğimiz hata payından bahsediyoruz.

Sample size calculations for a fixed power	Help
Power level between 0 and 1	0.08
Ratio of controls to cases	1.0

Genetic model free method	Help
Assume the Hardy Weinberg Equilibrium for cases?	no
Assume the Hardy Weinberg Equilibrium for controls?	yes

Gordon Heath Liu Ott error model	Help
ϵ_1	0.01
ϵ_2	0.01
Reset Next Page	

Soru: Genotipleme hatasını gerçekten göstermek mümkün mü?

Cevap: Mümkün. Genotipi bilinen kontroller kullanarak veriminizi test edebilirsiniz. Ayrıca, aile çalışmalarında, örneğin PedCheck gibi bir program kullanarak aile içi tutarlılığı test etmek ve bu sayede genotipleme hatalarını yakalamak mümkün. Bir de yüksek verimli sistemlerde hatalı olmasa da, genotipin belirlenemediği durumlar olabiliyor. Yani belirli bir sonuç alma "call rate" oranınız var. Bu oranı hatalı ya da belirlenemeyen sonuç olarak değerlendirmek mümkün.

Soru: Matematiksel model olmadan, bir altın standard genotipleme hatası düzeltme yöntemi var mı?

Cevap: Hayır. Burada sadece istatistik düzeltmeden bahsediyoruz. Amaç bireye ait doğru genotipi belirlemekse, farklı yöntemlerle incelemeyi tekrarlamakta yarar var.

Sonuçta genotipleme işleminde başarıyı değerlendirirken iki sorunla karşılaşmak mümkün:

1. Hatalı genotipleme

2. Genotipi belirleyememe, yani kullandığımız sistemin verimine dayalı hatalar. *Call rate* dediğimiz oran, genotipleme sonucunda bize ulaşan değerlendirilebilir sonuç oranıdır. Örneğin bir incelemede 100,000 genotipleme yaptığımızda, 99,000 başarılı sonuç alıyor, 1,000 sonuç belirsiz kalıyorsa, % 99 genotipleme başarısından (*call rate*) bahsederiz. Hatayı değerlendirmek için, bir başka genotipleme yöntemi ile sınamak ya da elimizdeki genotipi bilinen pozitif kontroller üzerinden girmek gerekir. Sonuç olarak, hatalı ya da eksik genotipleme oranı bizim son analizimizi yani ilişki ya da bağlantı tanımlamamızı ne kadar etkileyecek bunu belirlememiz ve matematiksel düzeltmeler yapmamız mümkün. Unutmamamız gereken burada amacımız tek bir bireye ait mutlak doğru genotipi bulmak değil, çalışma grubunun tamamının verileri ile gerçek bir genetik ilişkiyi (ya da bağlantıyı) tanımlamak.

Şimdi bir sonraki sayfaya geçelim. Bu sayfadaki veriler HWE bekleyip beklememeye

ya da genetik model olup olmamasına göre değişkenlik gösterebilecek alel ve genotip frekans verileri. İyi bir güç analizi yapmak istiyorsak, çalıştığımız genetik polimorfizmin sağlıklı kontrollerdeki alel frekansını bilmek gerekir. Ardından bizim hastalarımızda, sağlıklı kontrollerden ne kadar alel/genotip frekansı farkı öngördüğümüzü ya da ne kadar farkı yakalamayı hedeflediğimizi belirlememiz gerekiyor.

Örneğin sağlıklı kontroldeki minör alel frekansından % 10 farkı göstermeniz için gereken hasta sayısı ile % 20 farkı gösterebileceğiniz sayı büyük farklılık gösterecektir. Bu sayfada bana sorduğu sorular: genetik olarak modelden bağımsız ve HWE uyma zorunluluğu olmaksızın, hastalarda beklediğim 1/1 ve 2/2 genotipi, yani 1 ve 2 alelinin homozigot olma sıklıklarını, kontrollerde ise HWE uygun alel frekansını soruyor. Sağlıklı kontroller ile ilgili elimde veri yoksa bu konuda bilgi alabileceğiniz 2 internet adresi vereceğim: Bunlardan birincisi NCBI içerisinde yer alan dbSNP. PubMed programını hepimiz kullanıyoruz. Burada araştıracağınız anahtar kelimeyi (burada gen ismi) girdiğiniz satırın baş kısmında taramanın yapılacağı veritabanı yer alıyor. Burayı PubMed yerine SNP olarak değiştirip, tarama yapacağımız geni de yanındaki boşluğa yazdığımızda, o gendeki bilinen tüm tek nükleotid polimorfizmlerine ait listeyi görmemiz mümkün olacak. Bu verileri dilersem "limits" başlığını kullanarak sınırlandırabilirim. Sonuçları sadece insana ait veriler, aminoasit değiştiren ya da değiştirmeyen, kodlayıcı bölgede ya da intronlarda yer alan tek nükleotid polimorfizmleri (SNP) gibi sınırlandırmam da mümkün. Bunlar arasından seçtiğim polimorfizme girdiğim zaman, bu polimorfizme ait bilgiler ile farklı etnik gruplarda yapılmış genotiplemelerin sonuçlarını bulmak mümkün. Burada beyaz ırka, Afrikalılara, Asyalılara ait veriler bulunabilir. Bu gruplarda kaç kromozomun incelendiğini de görebiliriz. Türk popülasyon verileri genellikle Avrupalıların genetik yapılarına benziyor. Dolayısı ile bir benzetme yaparak alel frekansı tahmininde bulunabiliriz ve güç analizi için son sayfayı bu verilere dayanarak doldurabiliriz.

İlk kez çalışmak üzere bir SNP seçiyorsak,

daha fazla bilgi almak için poliformik bir örnek seçmeliyiz. Bazı SNP'ler çok düşük oranda polimorfizm gösteriyor, yani küçük alel frekansı çok düşük oluyor ve bizim bu polimorfizmden anlamlı sonuç alma şansımız oldukça düşük.

Sağlıklı kontrol verisine ulaşabileceğimiz bir diğer veritabanı HapMap projesine ait: <http://www.hapmap.org/index.html.en>. HapMap projesi uluslararası bir konsorsiyumun ürünü ve bugüne kadar tanımlanmış SNP'lerin nasıl haplotip oluşturduğunu ve bunların haritalamada/ilişki çalışmalarında nasıl kullanılabileceğini tanımlamayı hedefliyor. Bu projede 4 ayrı etnik gruba ait 270 kişinin verileri var. Kendi etnik grubumuzla ilgili veriler yok ve bu konudaki çalışmalar maalesef "Türk geni"ne ait bilgilerimiz yurtdışına kaçırılıyor diye yorumlanabiliyor. Oysa kendi etnik grubumuza ait polimorfizimlerin dağılımının bilinmesi ileride en büyük avantajımız olacak. Bu projeye Amerika'da yaşayan Avrupa kökenli beyaz 30 CEPH ailesinden ve Afrika kökenli Yoruba yerlisi (Ibadan, Nijerya) 30 aileden alınan trio örnekleri ile 45 Japon ve 45 Çinli (Han) örneği dahil edildi. Bu projeye ait internet sayfasında "*Browse project data*" bölümüne girdiğinizde önünüze

yeni proje sayfası açılacak. Buraya bir gen adı girdiğimizde, o gende taranan polimorfizmlerin 4 etnik gruplarındaki dağılımını grafik olarak da görmek mümkün. Beyaz ırkla Afrika, Japon ve Çin verilerinin çok benzediği polimorfizmler kadar çok farklı sonuçların olduğu polimorfizmler de göreceksiniz. Siz tecrübenize göre polimorfik olduğunu gördüğünüz ve fark gösterebileceğine inandığınız bölgeleri özellikle seçebilirsiniz. Ayrıca haplotip verilerini kullanabilirsiniz. Bu polimorfizmlerden herhangi birine tıkladığınızda, o polimorfizmin Avrupa kökenlilerde, Çinlilerde, Japonlarda ve Afrikalılarda alel ve genotip frekansını görme şansınız var.

Buradan yola çıkarak başladığımız hesaba geri dönecek olursak, diyelim ki hastalarda 1/1 genotip frekansı % 40, 2/2 frekansı % 10 ve sağlıklı kontrollerdeki 1 alleli frekansı % 40 olsun. Run PAWE dediğimiz zaman bize hata payı olmadan ve hatayı göz önüne alarak ne kadar hasta ve kontrol çalışmamız gerektiğini verecektir. Burada olduğu gibi hastalar ve kontroller arasında büyük bir fark öngörüldüğünde, genotipler arasında fark göstermek için 39 ya da 40 hasta ve kontrol yetebilecektir.

Genetic model free method without Hardy Weinberg Equilibrium	Help
case 11 genotype frequency	0.4
case 22 genotype frequency	0.1

Genetic model free method assuming Hardy Weinberg Equilibrium	Help
control 1 allele frequency	0.4
Reset Run Pawe	

Örneğimizde hastalarda 1 aleli frekansı % 65, kontrollerde % 40 olarak görülmektedir. Kontrollerdeki 1 alel sıklığı % 55'e artırıldığında genotipik farkı göstermek için gerekli örnek sayısı 2 kolda da 200 ve hata payı ile 209'a yükselmektedir. Alel düzeyindeki farkı % 5 yaptığımızda, alel düzeyinde farkı göstermek için çalışılması gereken örnek sayıları 736 (768), genotip düzeyindeki farkı göstermek için 589 (625) olmaktadır. Bugün incelenen polimorfizmlerden beklenen riskler (*odds ratio* OR) genellikle 2'den küçüktür ve bu düzeyde farkı göstermek için gerekli örnek sayısı oldukça büyüktür.

Katkı: O zaman bugün tanımlanmış olan bazı polimorfizm ilişkilerinde, mesela PTPN22 için, yeterli gücün olmadığını söyleyebiliriz.

Katkı: Güç hesabını yaparken bir deneye % 80 güç ile yapıldı demek, ben bu deneyi 10 defa yapıyor isem, 8 tanesinde aynı farklılığı görürüm mutlaka demek. Anlamı bu. Ama bir kere gördükten sonra güç hesabının bir anlamı kalmıyor. Bu konu yazarların artık söyledikleri bir laf değil. Eskiden söylenirdi bu, 20 sene önce söylenirdi. Ama şimdi söylenmiyor. Yani bir kere anlamlı çıktıktan sonra, artık gücü düşünmek gereksiz bir olay diye düşünülüyor. Güç başlangıçta çalışmayı planlarken önemli. Gücü eğer başta planlıyorsan önemli, sonradan hesaplamakla arada gerçekten bir fark var.

Bu söylediğiniz temelde pozitif bir ilişkiyi göstermek açısından anlamlı olabilir. Özellikle genetik ilişki çalışmalarına bakacak olursak, ciddi bir replikasyon, tekrarlanabilirlik sorunu görülüyor ve yeterli gücün olmamasına dayanıyor. Maalesef insanlar anlamlı bir şey gördükleri zaman çalışmanın gücünü unuttur hale geliyor. Oysa çalışma yeteri kadar güçlü değilse, az evvel simülasyon örneğinde yaptığımız gibi, siz çalışma sayısını artırdığınızda negatif sonuçları da görme şansınız artıyor.

Soru: PTPN22 geni ile ilgili çalışmalarda OR 1.5 ila 1.8 arasında veriliyor. Biz 1.6 bulduk. Ama Batı toplumlarından farklı olarak alel sıklığını % 20-30 değil, % 5 bulduk. Şimdi bizim bulduğumuz bu OR değerini nasıl yorumlamamız gerekiyor. Yani bu veri diğerlerinin yanında yer alabilir mi ve metaanalizlere katılabilir mi?

Cevap: Burada kritik olan sadece OR değeri değil, % 95 güven aralığının ne olduğu konusu. Dolayısı ile OR değerini, güven aralığını dikkate almadan değerlendirme şansımız yok. Bu da örneklem büyüklüğü ile çok ilişkili. Eğer küçük farklılıkların peşindeyse, yani sorduğunuz örnekteki gibi % 5'ten küçük bir farkı göstermek istiyorsak, 200 örnekte elde ettiğim sonuç 2,000 örnekle çalıştırdığımda farklı çıkacak. OR 2'den küçük olduğunda, gerekli örnek büyüklüğü inanılmaz artıyor. Bunu dikkate almadan, yani sadece bu farkı istatistiksel olarak bana OR=1.6 değerini veriyor dediğiniz zaman, yanlış payımız oldukça çok yüksek. Bu nedenle % 95 güven aralığı mutlaka dikkate alınmalı. Belki bu oran Avrupa'da bildirilen orana yakın. Ama bunu gösterebilmek için belki 500, 1000 ya da 1500 örnek çalışmak gerekli ki, bizdeki gerçek katkının ne olduğunu bilelim.

Bir başka önemli nokta da, değişik polimorfizmlerin hastalık patogenezeine sağladığı katkı oranlarının toplumlar arasında farklılıklar gösterebilmesi. Örneğin ankilozan spondilitte HLAB27'nin katkı oranı ülkemizde Kuzey Avrupa'dan daha düşük olabilir.

Soru: Bu konuşulanlardan sonra önce kontrol grubunu tarayıp, yani alel frekansını orada belirleyip ona göre karar verilebilir diye düşündüm.

Sağlıklı kontroldeki genotip dağılımını biliyor olmak çalışmayı planlarken çok önemli.

Bu nedenle dbSNP ve HapMap sayfalarını göstermek ihtiyacını duydum. Buradan bir fikir edinme şansımız var. Yine de bu ipuçları Türk toplumunu birebir yansıtan veriler değil.

Katkı: Bazı durumlarda veri bulmak da mümkün olmayabiliyor. Başka toplumlarda iyi tanımlanmış bir polimorfizmin alel sıklığı bizde sıfır olabiliyor.

Bunun tam tersi de geçerli. Örneğin Behçet hastalarında TNFRSF1A geninde R92Q mutasyon sıklığını % 5 civarında bulduk. Ama sağlıklı kontrollerimizde de % 5 civarında. Oysa Batı toplumunda bu sıklık % 1 veya altında. Dolayısıyla kendi sağlıklı kontrollerimizdeki dağılımı bilmeden anlamlılığı test etmek kolay değil.

Umarım uygulama yararlı olmuştur. Hepinize katılımınız ve katkılarınız için çok teşekkürler.

İTERAKTİF OTURUM II

Vedat Hamuryudan, Banu Çakır

Vedat Hamuryudan:Bu bölümde sizlere daha önceden dağıtılmış olan ve biri kontrollü ilaç çalışması, diğeri ise bir metaanaliz olan iki makaleyi inceleyeceğiz. Kontrollü klinik çalışmalar ve bu çalışmaların metaanalizleri hekimlerin günlük tedavi uygulamalarını etkilemede en belirleyici unsurlardır. Bu tür makaleler kanıta dayalı tıbbın hiyerarşik sınıflamasında epidemiyolojik veri kalitesi açısından en üstte bulunmalarına karşılık maalesef hepsi arzu edilen kalitede değildir. Bu durum, makalelerin kalite kontrolünü denetleyen yöntemlerin geliştirilmesine yol açmıştır. Bunlara bir örnek olarak Downs ve Black'in skrolama sistemini verebiliriz (1). Bu skrolama sisteminde 27 soru ile randomize veya randomize olmayan bir ilaç çalışmasının kalitesinin kontrol edilmesi amaçlanıyor. Örneğin, çalışmanın amacı, net olarak ifade edilmiş mi? Çalışmaya alınan hastaların özellikleri açıkça belirtilmiş mi? Randomizasyonun ne şekilde yapıldığı açıkça anlaşılıyor mu? Makaleyi inceleyerek bu sorulara 0 veya 1 puan verilir. Bizim inceleyeceğimiz kontrollü çalışmaya bu skrolamayı uyguladığımızda 27 üzerinden 18 puan aldığını gördüm ki, bu orta kalitede bir çalışma anlamına geliyor. Daha basit bir başka skrolama sistemi var ve JADAD olarak adlandırılıyor (2). Beş soru içeren bu yöntemde puanlama şöyle oluyor: Çalışma randomize olarak tanımlanmış yani random ya da rastgele kelimeleri kullanılmışsa 1 puan, yoksa 0 puan; randomizasyon yöntemi tanımlanmış ve yeterli ise 1 puan, yetersiz ise 0 puan; çalışma çift kör olarak tanımlanmış ise 1 puan; çift körleme yöntemi tanımlanmış ve yeterli ise 1 puan, yetersiz ise 0 puan; çalışmadan çıkarılma veya ayrılma açıklanmış ise 1 puan verilmekte.

İnceleyeceğimiz kontrollü ilaç çalışmasının amacı, rebamipide adlı ilacın mevcut tedaviye ek olarak kullanıldığında Behçet hastalığının ağız yaralarına olan etkisini göstermek (3). Başlığında çift kör, plasebo kontrollü olduğu belirtilmiş. JADAD skrolamasına göre hesaplandığında 5 üzerinden 4 puan almakta. Rebamipide, Türkiye'de bulunmayan, mide mukozasını kaplayarak etki eden gastrik mukoprotektif bir ilaç. Makale, Behçet hastalığında oral ülserlerin en yaygın şikâyet nedeni olduğunu, hastaların yaşam kalitesini düşürdüğünü ve tedavilerinin çok etkili

*İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi, İstanbul
Hacettepe Üniversitesi Tıp Fakültesi, Ankara*

olmadığını belirten, yani sorunu ortaya koyan cümleler ile başlıyor. Ardından ilkin Rebamipide'i tanıttıktan sonra, daha önce yapılmış kontrolsüz çalışmalarda bu ilacın etkili bulunduğunu bildirerek 3, 4 ve 5 numaralı kaynaklara gönderme yapıyor. Aslında bu giriş oldukça iyi ancak, "bu çalışmalar nelerdir?" diye bahsedilen kaynaklara baktığınızda, sadece 3 numaralı kaynağın aynı grup tarafından yapılmış bir kontrolsüz çalışma olduğunu, diğer iki kaynağın ise Behçet hastalığını anlatan derlemeler olduğunu görüyoruz. Ortada sadece tek ve o da kendileri tarafından yapılmış çalışma varken, bunu çalışmalar diye vermek benim görüşüme göre makalenin güvenilirliğini bir anda zayıflatan bir yanıştır. Makale daha sonra Rebamipide'in aktif oksijen inhibe edici özelliği olduğunu ve bu nedenle aftlara iyi gelmesi gerektiğini ileri süren bir cümle ile devam ediyor. Bu sonuca nasıl ulaşıldığı belli değil. Daha sonra aftlar yaşam kalitesini düşürdüğünden, bu ilacın hastaların yaşam kalitesini yükseltmesinin beklendiğini ve bu nedenle bu kontrollü çalışmanın planlandığı belirtiliyor. Yöntem kısmında çalışmanın 1994 – 1996 yılları arasında 6 merkezde toplam 35 hasta ile gerçekleştirildiğini öğreniyoruz. Bir aylık gözlem ve 6 aylık tedavi süresi var. Behçet hastalığının tanısı Japon kriterlerine göre konulmuş. "Son 1 ay içinde en az 1 hafta süre ile ağız yarası olan hastalar" çalışmaya alınmış. Hastaların son 1 ay içinde başka ülser ilacı kullanıyor olmaları tek dışlama kriteri. Hastalar Behçet hastalığı nedeni ile kullandıkları ilaçlara devam etmişler, ancak doz değiştirmeye veya yeni ilaca başlamaya izin verilmemiş. Bu aşamada görüş belirtmek isteyen var mı?

Hasan Yazıcı: Bu çalışma Ağustos 1994 ile Aralık 1996 arasında yapılmış diye belirtmek güzel, ancak ne zaman yayımlanmış diye baktığımızda 2003'te yayımlandığını görüyoruz. Derginin kabul süresi çok uzun, onun için desek yine bu kadar uzun süre mantıklı olmuyor. Çalışmanın bitirilmesinden yayımlanmasına kadar geçen 7 senelik süre alışılmış bir durum değil.

Yabancı bir dinleyici: Hastaları toplamak için ne kadar süre geçtiğini merak ediyorum. Burada gerçekten de 35 hasta için 2 yıldan fazla süre geçmiş. Hastalık mı nadir görülüyor? Yoksa hastaların toparlanmasındaki sorun neydi?

Hasan Yazıcı: Japonya'da bu hastalık pek nadir değildir ama uygun hasta bulmak pek kolay değildir sanırım. Hasta sayısı beni pek rahatsız etmiyor; beni burada rahatsız eden, yayımlamak için geçen 7 yıl zaman. Neden daha önce yayımlanmamış, cevabı bilmiyorum. Benim önerim her zaman çalışmanın yapıldığı, analizin yapıldığı, kabul edildiği ve yayımlandığı tarihe bakmaktır.

Vedat Hamuryudan: Artık kontrollü çalışmaların yayımlanabilmeleri için daha başlamadan önce belirli veri tabanlarına kayıt edilmeleri gerekiyor.

Hasan Yazıcı: Bir de unutmamak lazım ki, bu bir Japon çalışması ve Japonlar kendi dillerinde yazmayı tercih edebiliyorlar, dolayısıyla belki Japonca çıkmıştır, onu kontrol etmek gerekir. Ancak PubMed'e giren bir çok Japon dergisi var ve onlarda yayımlayabilirlerdi.

Vedat Hamuryudan: Randomizasyonun yapılışı, "*by using a random allocation list*" şeklinde tanımlanmış. Ancak, nasıl "*random allocation list*" yapıldı? Altı merkezli çalışma olduğunu biliyoruz. Her merkez hasta aldıkça bir yere bildirdi, orası da listeye göre plasebo veya ilaç mı gönderdi? Yoksa randomizasyon bloklar halinde mi yapıldı? En baştan 35 hasta ile yapılacağı belli miydi? Bu çalışmanın gücü çok düşük olsa gerek, ancak makale başında bu bir öncü çalışmadır diye belirtiliyor. Sonuç olarak randomizasyon nasıl yapıldığı esasında çok belli değil. Günde 3 defa aktif ilaç veya ona eş değer gözüken plasebo uygulanıyor. Hastalar ilk önce bir ay süre ile gözleniyor ve daha sonra 6 ay boyunca tedavi alıyorlar. Her ay vizitte ilacı kullanıp kullanmadıkları sorgulanıyor. Behçet hastalığının oral ülserlerinin tedavisi ile ilgili olarak bizim de yaptığımız çalışmalar var. Değerlendirme o kadar kolay değil. Hastaya tedaviyi verip onbeş gün veya bir ay sonra kontrole çağırdığınızda nasıl değerlendireceksiniz? Hastada o an mevcut ülserleri saymak bir yöntem, ancak o zaman arada olup iyileşen ülserleri atlamış oluyorsunuz. Vizitleri haftalık yapabilirsiniz ; ama o zaman bir önceki hafta olup da iyileşmemiş ülserleri nasıl ekarte edeceksiniz? Hastaya her yeni çıkan ülseri kaydetmesi için günlük tutmasını söylemek mümkün. Ama ülkemiz koşullarında bunun da pek yürümediğini daha önceki

tecrübelerimizde gördük. Bu çalışmada da hastalara günlük tutarak her gün mevcut olan ülserleri kaydetmeleri söylenmiş. Diyelim bugün hastanın ağzında 5 yara var, 5 diye yazacak. Ertesi günü hiçbiri iyileşmedi ise yine 5 diye yazılacak, bir sonraki gün biri iyileşti ama iki tane çıktı ise 6 yazılacak. Bu çalışmada aftlar aylık toplanarak bir total skora ulaşılmış. Bir de aynı günlük ülser sayısı gibi, günlük ağrı skorlaması yapılmış. Daha sonra da ziyaretlerde hekimin global değerlendirme yaptıkları yazılı ancak bunun nasıl yapıldığı açıklanmamış. Ek olarak 12 tane hematolojik, 13 tane biyokimyasal test yapıldığı ve bunlardaki en ufak anormalleşmenin kötüleşme olarak kabul edildiği belirtiliyor. “Bu testlerin ölçülen oral ülserlerle ilişkileri nedir?”, “niye bakıldı?” belli olmadığı gibi, bu kadar çok sayıda test yapınca anormal sonuç bulma olasılığınız da artacaktır. Analizler tedaviye kalkışma yani “*intention to treat*” yöntemine göre yapılmış. Sonuçlar ortalama ve standart hata, yani “*standard error of mean*”, olarak verilmiş, genel olarak kullanılan standart sapma yani “*standard deviation*” kullanılmamış.

“*American Journal of Medicine*” isimli önde gelen genel dahiliye dergisini hepimiz bilirsiniz. Bu dergiye makale gönderilirken uyulmasını istediği şartlardan bir tanesi sonuçların standart hata değil, standard sapma olarak verilmesi. Standart hata, standart sapmanın karekökü ve standart sapma büyük olduğu zaman yani dağılımın homojen olmadığı durumlarda kullanılıyor.

Banu Çakır: Standart hata ve standart sapma çok karışır gerçekten. Bunun, dediğiniz gibi bazen bilinçli olarak da yapılabileceğini düşünüyorum. Tabii amacın ne olduğunu söylerlerse ve kullanımı uygunsa, yanlış olmayabilir. Ama burada amaç çok açık değil. Standart sapma, dağılım/yayılm ile ilgili bir ölçüt. Dağılımın normal, çan eğrisi şeklinde, simetrik olduğu durumlarda kullanımı en uygunu standart sapma; toplumdaki kişilere ait değerler arasında ortalama ya kıyasla ne kadar bir farklılık olduğunun bir ölçütü: gözlemlerin normal dağıldığı bir toplumda, toplum fertlerinin % 95’ine ait bireysel değerlerin, toplumdan alınan örnekten elde edilen ortalama değer 2 standart sapma altı ve 2 standart sapma üzerindeki değerler

arasındaki dilimde yer aldığı kabul edilir. Standart hata ise, hesaplayabileceğiniz ortalamalara ait bir sapma. Yani, aynı toplumdaki, benzer sayıda, farklı örnekler alınsa ve örneklerle ait ortalamalar hesaplasa idi, bu ortalama değerlerinin, tek çalışmadan elde edilecek ortalama ya kıyasla, sayısal eksende ne kadar sağda veya solda olabileceğini söylüyor. Başka bir ifade ile, bizim, bir çalışmadan elde ettiğimiz ortalama değerden yola çıkarak, topluma ait ortalamanın hangi değer aralığında yer aldığı konusunda bir fikir veriyor. Topluma ait ortalama değerinin, % 95 olasılıkla, çalışmaya ait ortalama ± 2 standart hata aralığında yer aldığı kabul edilir.

Bu çalışmada amacımız başlangıç karakterleri açısından grubu tanımlamak olduğu için, standart sapmayı vermek uygun olur. Ortalama ve ± 2 standart sapma aralığını verdiğimiz zaman; grubun % 95’ine ait bireysel değerlerin yayıldığı aralığı tanımlayabiliyoruz. Şimdi, ikinci durumda, aft sayısına bakarken, standart hata kullanımı doğru olabilirdi: “Ben size ortalama aft sayısını söylüyorum, ama bu aft sayısı ortalama toplumdaki buna eşit olmayabilir, hatasını da söylüyorum, dolayısıyla toplumdaki gerçek aft sayısı ortalama bu aralıktadır” derlerse, o zaman doğru. Ama burada bunu dağılım için kullanmaları doğru değil.

Önemli bir özellik daha var. Belki şu aşamada gündeme gelmiyor, daha çok randomize klinik çalışmalardan bahsettiğimiz için. Bir grubu temsil etmek, toplumu temsil etmek söz konusu olduğu zaman örnek üzerinden elde edilen değerden yola çıkarak (örneğin, artrit sıklığı) toplumu temsil etmek söz konusu olduğu zaman (örneğin, toplumdaki artrit prevalansı), o zaman standart hatayı kullanarak, topluma ait değeri yansıtmak için güven aralığını hesaplıyoruz ki, bu durumda standart sapma değil, standart hata kullanımı doğru olur.

Burada, dediğiniz gibi, standart sapma vermek daha doğru olurdu. Çalışma başlangıcında yapıldığı gibi.

Bir de, daha önce söylediğiniz bir konuyu ben de tekrar vurgulamak istiyorum. Şimdi, baştan bir ölçüm yapıyorlar: aftları tek tek sayıyorlar. Aylık sayımlar yapıyorlar. Ondan sonra hiç daha önce açıklamaları yapılmamış ifadeler ile, “orta”, “ağır”, “ciddi”, “hafif” diyorlar.

Yine bilmediğimiz bir şekilde, “ilerleme var” diyorlar, “gerileme var” diyorlar. Bu çok yanlış bir şey. Sayısal bir tabanda karşılaştırma yapmıyorlar. Başlangıçta saymışsak emeğe de yazık diye düşünüyorum. Onun için böyle uygulamaların uygun olmadığını tekrar vurgulamak istiyorum.

Vedat Hamuryudan: Sonuçlara baktığınızda aft sayılarının oldukça yüksek ve hastadan hastaya çok değişken olduğunu görüyoruz.

Bir dinleyici: Siz biliyorsunuz ama okumayan arkadaşlar için de tekrar hatırlatmakta yarar var. Bunların hepsi yeni çıkan değil, devam eden aftlar da var. Bu da çalışmayı değerlendirmeyi kötü yönde etkiliyor. Çünkü olanlar ile yeni çıkanları ayırt etmiyor.

Bir dinleyici: Hastalık tanısı Japon kriterlerine göre konmuş idi. Gastrointestinal tutulumu olanlar var ve onlarda aft sıklığı daha fazla. Burada heterojen bir grup söz konusu. Kiminde göz tutulumu, kiminde gastrointestinal tutulum ön planda. Çalışma daha homojen bir hasta grubunda örneğin sadece mükokütanöz tutulumu olanlar arasında yapılabilirdi. Ayrıca kullanılan diğer ilaçlar konusu var. Örneğin, hastaların önemli bir kısmı kolşisin alıyor. Bunun sonuca etkisi belli değil. Hastaların birincil değerlendirme ölçütlerinden bir tanesi hekimin global değerlendirmesi. Plasebo grubunda hekimler % 36 oranında, yani 14 hastadan 5’inde iyi sonuç aldık diyorlar. Rebamipide grubunda ise, bu oran % 65 olarak veriliyor. Burada ki-kare testi uygulandığında anlamlı bir fark olmadığını görüyoruz. Bunun dışında bütün çalışma süresince görülen aft sayılarının aylara göre ortalama dağılımları bir grafikte veriliyor ve iki grup karşılaştırılıyor. Plasebo grubunda tedavi öncesi gözlem döneminde azalma, rebamipide grubunda artış var; tedavi ile birlikte rebamipide grubunda aft sayısı baskılanmış giderken, son ayda artış görülüyor. Ancak bu grafikte hasta sayılarının farklı olduğunu görüyorsunuz. Onsekiz hasta ile başlanıyor, sonra 16, 14 hasta ile devam ediliyor. Diğer grupta da 14 ile başlanıp 10 ile bitiriliyor. Aylara göre fark var mı, bunların sapmaları nedir, verilmemiş.

Hasan Yazıcı: Japonlar pek kontrollü ilaç çalışması yapan bir ulus değil.

Behçet hastalığında daha önce yaptıkları tek kontrollü ilaç çalışması siklosporinle yapılmıştır. Bahsettiğim çalışmada kolşisine cevap vermeyenlere siklosporin vermişlerdir. Kontrollü ilaç çalışması Japon geleneği değil. İlacın iyi geldiğini söylüyor, ancak hakkını veriyor ve plasebo grubu tam anlamı ile bir plasebo grubu değildi çünkü ek olarak aktif ilaç alıyordu diyor. Unutmayın % 63 bir grupta, % 33 bir grupta kolşisin kullanımı vardı. Ayrıca diyorlar ki, bizim pilot çalışma uygun bir gruptur, küçük değildir; ancak bunu neye göre söylüyorlar, belli değil.

Vedat Hamuryudan: Rebamipide, doku tamir mekanizmalarına etkili olduğu için ülserler muhtemelen daha derinleşmeden iyileşti diye ileri sürüyorlar. Bir başka yerde pomad gibi lokal tedavilerin uygulanma zorluğu var; “biz günde 3 kere rebamipide kullanarak lokal tedavilerin bu zorluğunun üstesinden geldik” diye belirtiyorlar. Bir başka yerde de, “madem bu ilaç afta iyi geliyor o zaman Behçet hastalığının diğer semptomlarına da iyi gelmesini bekleriz” diye bir hipotez ileri sürüyorlar. Son olarak, daha kalabalık hasta grubuna ihtiyaç var diyerek bitiriyorlar.

Hasan Yazıcı: Şimdi Behçet hastalığında oral aftlar üzerine bir çalışma planlıyor olsanız mevcut Helsinki Deklarasyonu kurallarına göre plaseboya karşı bir çalışma yapmak doğru mudur, değil midir?

Haner Direskeneli: Plaseboya karşı yapabilirim çünkü burada, romatoid artritteki gibi tedavide geç kalma sonucu ileride daha büyük sorunlara yol açacak etik problem söz konusu değil. Tek problem hastanın ağrısının geçmesine izin verilmemesi gibi bir şey olabilir ama o zaman da bir kurtarıcı “rescue” ilaç kullanımını protokole koymak mümkün.

Yabancı bir dinleyici: Kısa süreli çalışma yapmanın etik olup olmadığını tartışıyoruz. Plasebo kontrol olmadan herhangi bir ilaç için noninferiority çalışmasına diğer gelen ilaçlarla geçmeden, örneğin romatoid artrit çalışmalarında, uygun etkinlik verilerine sahip olabilir miyiz? Ne düşünüyorsunuz? Romatoid artritte her zaman metotreksat kullanırız ve her şey de ne olursa olsun, onunla karşılaştırılır.

Hasan Yazıcı: Romatoid artritte gerçekten iyi bir faz III çalışması yapıyorsak, metotreksatin

kullanılması gerektiğini düşünüyorum. Ben FDA gibi düşünmüyorum; önce ilacın plaseboya göre etkinliğinin gösterilmesi gerektiğine katılmıyorum. Çünkü hastalardan *informed consent* almak da çok zor artık. Tüm değişiklikler düşünüldüğünde, Behçet hastalarında plasebo kontrollü çalışma yapabilirim.

Örneğin, sadece % 30'da ACR 20 cevabınız varsa bunlar plaseboya benzerdir ve yapmazsanız, sadece % 30 cevaplı yeni bir monoklonal antikor içeren ilaç, iki kollu bir çalışmada metotreksata göre kötü değil (non-inferior) olarak çıkabilir. Genel koşullarda, düşük cevap oranı ile gerçek ilaç ile plasebo arasında bir yerlerde, hala daha standart ilaca göre noninferior kalabilir ve belki de sadece iki hafta için yine de plasebo kontrollü uygun bir çalışmayı gerektirebilir.

Banu Çakır: Oturumun ikinci yarısında tartışmak üzere bir metaanaliz seçtik (5). İki sebeple: birincisi, geçen sene, Dr. Ioannidis'in bir metaanaliz konuşması vardı; bu yıl buna devam olsun istedik, toplantılar arasında birliktelik yakalamak için. Onun dışında, bu metaanalizde sadece randomize kontrollü çalışmalar değil, aynı zamanda vaka kontrol ve kohort tipi çalışmaların da metaanalizi yapılmış ve birlikte sunulmuş. Nonsteroidal antiinflatuar ilaçlar sizlerin yaygın kullandığınız bir ilaç grubu ve miyokard infarktüsü ile olan ilişkisi halen gündemde bir konu.

Öncelikle ne yapmışlar diye özetleyerek başlamak istedim. Çalışmadaki amaç, geleneksel nonsteroidal antiinflatuar ilaçlar ile COX spesifik nonsteroidal antiinflatuarları, miyokard infarktüsü ile olan ilişkileri açısından karşılaştırmak. Bu amaçla üç tane sistematik derleme yapmışlar: bunlardan biri vaka kontrol çalışmaları, biri kohort çalışmaları, bir diğeri de randomize kontrollü klinik çalışmaları toplamaya yönelik. 1996 – 2000 yılları arasında dergilerde yer almış ve farklı veri tabanlarında yayımlanmış tüm eserlere ulaşılmaya çalışılmış; sadece İngilizce yayınlar ve en az 100 kişilik çalışmalar alınmış; bu çalışmaların ham verilerine ulaşabilmek esas alınmış. Bu konuda, varsa, yayımlanmış metaanalizleri de almışlar. Bunların üzerinden tartışmada gideceğiz, bu nedenle şu aşamada çok bahsetmek istemiyorum. Ama önemli özellik; çalışmalar gerçekten çok

heterojen ve bu konuyu tartışırken gündeme tekrar gelecektir. Birleştirmeleri çok doğru olmayan çalışmalar da var derleme içinde. Bunu kendileri de belirtiyor ve değişik kontroller yapmaya çalışıyorlar: örnek büyüklüğüne göre, çalışmanın yapıldığı ülkeye göre, maddi bir destek alınıp alınmadığına göre, çalışmanın tipine göre, çalışmaya alınma kriterlerine göre ne tür kontroller yaptıklarını açıklıyorlar.

Ben metaanalizler konusunda geçen yıl neler konuştuk, kısaca hatırlatmak istiyorum. Metaanalizler epidemiyolojik veri kalitesi açısından en değerli çalışmalar. Farklı ülkelerde farklı isimler ile adlandırılabilir. Metaanalizlerin popüler olmasının en önemli nedeni tek tek çalışmaların örnek büyüklüğünün yeterince büyük olmaması. Çalışmanın gücü bazen çalışma öncesinde hesaplanmamış oluyor, sonucu kestirmek için yeterli gelmiyor, benzer çalışmaları birleştirmek uygun oluyor veya bazen “şüpheli” ilişki bulunan çalışmalar bir araya getirilip, veriler harmanlandığında merak edilen ilişkiyi göstermek için çalışma gücü artırılıyor. Ayrıca, metaanaliz yaparak, bu çalışmalar gerçekten heterojen mi değil mi, anlaşılabilir. Çalışmaları tek tek okuduğumuzda sonuçlarından etkileniyoruz ama bir araya getirdiğimizde “acaba aynı şeyi mi söylüyor, yoksa farklı şeyler mi, hangi aşamalarda farklı şeyler söylüyor?” Bunu görmek açısından, metaanalizler önem taşıyor. Metaanalizler tabii bir de, çalışmalarda ne tür hatalar yapıyoruz, görebilmek açısından önemli. Çalışmalar birlikte değerlendirildiğinde, özellikle de yayımlanmamış çalışmalar da incelenebilmiş ise; “*publication bias*” var mı? Belki biraz önceki çalışmada da söz konusu olabileceği gibi, “*time lag*” var mı? Yani çalışma yapılıyor fakat çalışma yapılırken ya da basıma dek o kadar uzun süre geçiyor ki bazı bulgular geçerli olmayabiliyor. Bu tür hataları saptamak mümkün oluyor.

Metaanalizlerde bulguların kalitesinden bahsederken unutmamız gereken konu, bir metaanalizin ancak derlediği çalışmaların veri kalitesi nispetinde kanıt sağlayabileceği.

Hasan Yazıcı: Hiyerarşide metaanalizden daha önemli bir şeyin olması da gerekli. 5 – 7 tane gücü yetersiz kontrollü çalışmayı bir araya getirip sonuç çıkartmak yerine iki tane birbirini

tekrarlayan güzel kontrollü çalışma bana daha önemli geliyor.

Banu Çakır: Kümülatif metaanalizlerde araştırma sonuçlarını yapıldıkları yıllara göre ekleye ekleye gidiyoruz, aynı sebeple. Bir süre sonra metaanalizler de geçerliliğini kaybediyor. Cochrane Grubu tarafından yapılan metaanalizlerde bu tür bir güncelleme ile yıllar içinde eklenen çalışma verileri ile kanıtın (varsa özelliklerindeki değişikliklerin) takibi amaçlanıyor.

Metaanalizleri hepimiz okuyoruz ve sonuçlarından yararlanıyoruz. Bugün birlikte değerlendireceğimiz bu metaanaliz için özel bir şey yapmak istedim. Bazı tartışma soruları hazırladım. Bu soruları yanıtlayarak, metaanalizin üzerinden birlikte gitmek istiyorum. Bu şekilde, bir metaanaliz okurken yöntemsel açıdan nelere dikkat etmeliyiz, onu da tartışabilelim istedim. Sizlerle bu sorular üzerinden gitmek istiyorum. Bu çalışmada gözlemsel araştırmalar ile randomize klinik çalışmalar bir arada bildiriliyor. Neden? Çünkü bunlar aslında birbirlerine alternatif değil. Yani biri diğerine üstündür demek her zaman doğru değil. Farklı tipte araştırmaların kendi içerisinde avantajları ve dezavantajları var. Randomize klinik çalışmalar, *efficiency* (laboratuvar şartlarındaki etkililik) ölçen çalışmalar: Her şey mükemmel ve planlanan şekilde uygulandığında ortaya çıkan etkili ölçebiliyor. Gerçek yaşam ise laboratuvara çok uymayabiliyor. Özellikle bizim gibi genetik açıdan çok heterojen toplumlarda elde edilen sonuçlar (etki, advers etki vb.) her zaman klinik çalışmalarda incelenmiş gruba uy(a)madığı için farklı sonuç elde edebiliyoruz. Uzun latent dönemi olan, nadir görülen advers etkiler de, benzer şekilde, az sayıda kişiyi inceleyebilen, kısa izlem süreli randomize klinik araştırmalarda ortaya çıkmayabilir. Bu nedenle, farmakovijilansa yönelik çalışmalar, güvenlik çalışmaları, ağırlıklı olarak gözlemsel tipteki araştırmalara gereksinim göstermektedir. Bunun dışında, toplumlarda kültür, ekonomik durum, doktora ulaşım, hastalığı ve/veya doktoru kabul etmek, ilaca uyum (compliance) gibi özellikler de ilacın etkisini değiştirebiliyor. Bu metaanalizde de gözlemlendiği gibi, gözlemsel çalışmalardan gelen sonuçlar ile randomize klinik çalışmalardan gelen veriler desteklenir özellikte ise, bu durumda “nedensel” (*causal*) ilişkiler konusundaki

kanıtlarımız güçlü kazanıyor. Ben bu metaanalizi okurken bazen rahatsızlık duydum, çalışma tiplerini tanımlarken herkes farklı tanımlar kullanabiliyor; benim çok alışkın olmadığım tanımlar var. “*Cohort as case control*”, “*case control as cohort*” benzeri tanımlar kullanılmış. Bu çalışmaların orijinallerinin nasıl yapıldığını inceledim. Bu çalışmalar, nested vaka kontrol tipli çalışmalar: yuvalandırılmış vaka kontrol çalışmalar olarak kullanıyoruz Türkçe’de. Bunların özellikleri ne?: kohort çalışma tabanlarında oturtulması, eş zamanlı olarak aynı kohorttan farklı nested vaka kontrol çalışma yapılabilir; kohort üyelerinin sadece sonuçlarının takibi yeterli olabilir; kohorta göre daha kolay ve ucuz çalışmalar; ama aynı zamanda risk ve rölatif riskler hesapları yapılabilmesi açısından son derece yararlı. Bu tür melez çalışma tasarımları kullanıldığı durumlarda ek ölçütler hesaplayabilmek veya kanıtın kalitesini artırabilmek mümkün. Bir önceki çalışmada Vedat Bey söyledi, JADAD kullanmışlar. Buradaki randomize klinik çalışmaları değerlendirirken CONSORT, benzer şekilde gözlemsel araştırmalar için de STROBE kullanılmış. STROBE 2004’ten beri gündemde, henüz yaygınlık kazanmış diyemem, ama 2003 yılların sonu 2004 başı gibi, gözlemsel araştırmalarda kanıtın kalitesi kontrol edilemiyor diye bir standardizasyon geliştirilmesi istendi. Hatta başında, ismi “*observation* çalışmaların raporlanmasında standartlaşma” diye geliştirilmişti, sonra beğenilmeyince STROBE kullanılmaya başladı. Vedat hocanın biraz önce diğer çalışma için söylediği şeyi burada da yapmışlar. Ancak, buradan aldıkları skor her zaman bu çalışmanın iyi olduğunu, güvenilir olduğunu göstermiyor, bu örnekte de bu sıkıntımız var. O nedenle, çalışmaların ham verilerini elde etmeye çalışmışlar. Ham veri ile neler yaptıklarını yazmışlar. Bu güzel. Ham verileri kullanarak incelenen ilişkiler için ilişkinin gücü ve standart hatasını kendileri tekrar hesaplamışlar.

Bu metaanalizin sonuçlarına dayalı kararlarımızı, “gerçek”ten çarpıtacak tarzda etkileyebilecek sistematik hata kaynakları neler olabilir diye baktığımızda, ilk dikkati çeken özellik, kanıt kalitesi farklı çalışma tiplerini birlikte değerlendiriyor olması. Değerlendirilen çalışmaların tasarımına özel sayısız kısıtlılık ve taraf tutma gözlenebilir. Öte yandan, farklı

çalışma tiplerini, kendi özgün avantaj ve dezavantajlarını göz önüne alarak değerlendirme yapabilirsek, bu bir hata kaynağı olmaktan ziyade, kanıtın gücünü arttırıcı bir özellik olarak da kullanılabilir. Birleştirilen çalışmaların özellik ve sonuçlarının heterojen olması bir “özet” sonuca ulaşmada sorun yaratabilir; bunu düzeltmek için alt grup analizleri (tabakalı analizler) yapılmış. Çok sayıda alt grup analizi yapılması söz konusu olduğu durumlarda, örnek hacminin bu tür analizlere izin verecek büyüklükte olup olmadığına dikkat etmeliyiz. Olası sistematik hata kaynakları arasında, sadece İngilizce dilinde yazılmış makale ve derlemelerin alınmış olması, farklı dilde yazılmış makale ve derlemelerde bulunan NSAİ ilaç-Mİ ilişkisinin burada incelenenden farklı olduğu ölçüde, bu uygulama taraf tutmaya neden olabilir. Benzer şekilde, yayıncıların sık olarak “fark bulan” araştırmaları yayımlama eğiliminin, yayımlanmış eserler üzerinden yapılan metaanalizlerde fark bulma lehine bir taraf tutmaya neden olabileceği de unutulmamalıdır. Çalışmaya katılması için teklifte bulunan bazı hastaların çalışmaya katılmaması; çalışma grubundaki kişilerin çalışmayı, bitmeden terk etmesi; erken dönemde ölen hastaların vaka olarak gruba alınamaması; hastaların ilacı önerilen doz ve saatte almaması; hastaların /kişilerin son durumlarının uygun ve yeterli şekilde tespit edilememesi ve olası karıştırıcı faktörler için analizlerde kontrol sağlanamaması gibi bir çok sistematik hata kaynağı bu metaanalize katılan çalışmaların bir ya da daha fazlasında yer almış olabilir ve bu da sonuçların gerçekten farklı yorumlanmasına yol açabilir. Çalışmada bu hata kaynaklarının bazılarının değinildiği ve etkilerinin tartışıldığını görüyoruz.

Metaanalizlerde, bir araya getirilen çalışmaların özelliklerinin benzerlik göstermesi önemlidir. Çalışmaların planlanması ve uygulanması ne kadar benzer ise, bir araya getirildiğinde alınacak sonuçların gerçeği yansıtırma gücü ve güvenilirliği o denli iyi olacaktır. Çalışmaların birbirinden farklı olması durumunda, “heterojenlik”ten bahsedilir. Bu çalışmada da hem klinik hem de istatistiksel açıdan heterojenlik değerlendirilmesi önemlidir. Çalışmada bu durumun değerlendirilmesine yönelik çabalar olduğu görülüyor. “Klinik” heterojenliği yok etmemiz çok mümkün değil,

ama bunu metaanaliz içinde tarif etmemiz önemli; tabii olası etkilerinin boyut ve yönü hakkındaki görüşlerimiz ile birlikte. İncelediğimiz metaanalizde “İstatistiksel” heterojenliği değerlendirmeye yönelik bazı testler yapıldığı ve tablolarda rastgele etkiler modeline göre hesaplamalara da yer verildiği izleniyor. Çalışma bulguların istatistiksel olarak birbirinden anlamlı fark göstermediği durumlarda, başka bir ifade ile bu çalışmada *odds* oranı değerlerinin, incelemeye alınan tüm vaka kontrol çalışmalarında “benzer (aynı)” olduğu varsayımı yapılabiliyor ise, “sabit etki modeli” kullanılabilir. Öte yandan, eğer sistematik derleme içinde incelenen farklı çalışmalar benzer bir etkiye işaret etmiyor, ancak, nokta değerler belli bir orta değere yakın, tanımlanabilen bir dağılıma uyuyor ise, “rastgele etkiler modeli” kullanılması daha uygun olur. Genel olarak rastgele modele alt etki gücü değerleri daha konservatif (tip I hata riski daha az) kabul edilir, ancak, yakın zamanda rastgele modellerden elde edilen güven aralığı değerlerinin sabit modellerden elde edilen değerlerin güven aralıklarından her zaman daha geniş olmadığı konusunda da makale ve tartışmalara rastlıyoruz. Genel olarak, en uygun kabul edilen yöntem, sabit ve rastgele etki modellerinden elde edilen nokta değer ve güven aralıklarını inceleyerek, aynı ise fark etmez ama, farklı ise daha konservatif olanı esas almak. Sanırım bu sebeple, bu makalede ilişkilere ait etki gücü değerleri verilirken, hem sabit hem de rastgele modele ait hesaplar, birlikte sunulmuş. Ayrıca, metaanalizde sadece nokta değerin ele alınması doğru değil. Bu nokta değerden olan sapmayı değerlendirmek, başka bir ifade ile çalışmaların birbirinden ne kadar farklı olduğunu göstermek; eldeki veriler imkan verdiğinde “geniş perspektiften” kanıta bakmak da metaanalizlerin önemli amaçlarından biri şüphesiz. Heterojenlik çok belirgin olduğu durumlarda tek bir özet değer (RR/OR) verilmemesi de tercih edilebilir. Bu çalışmada, heterojenlik için istatistiksel testler ve alt grup analizleri yapıldığı görülüyor.

Makalenin vaka kontrol tipli çalışmalara alt metaanaliz kısmında Peto’nun *odds* oranı değerinin hesaplandığı görülüyor. *Peto’nun odds* oranı, incelenen olayın nadir olduğu ve tedavi etkisinin düşük olduğu ya da vaka kontrol gruplarındaki kişi sayısının benzer olduğu

durumlarda kullanılması uygun bir değerdir ve gruba ait ortak değeri yansıtır, bu çalışmada da yerinde kullanıldığını görüyoruz. Makalede çalışmaların birbirinden farklı özellikleri için bir takım alt grup analizleri yapıldığı belirtiliyor demıştim. Alt grup analizlerinin hangi durumlarda yapılmasının uygun ve yerinde olduğuna dikkat etmeliyiz. Alt grup analizleri, tek bir çalışma dahilinde de yapılabileceği gibi, bu tür metaanalizlerde de kullanılabilir. Çalışmada alt grup analizlerinin bir çoğunun NSAİ ilaç kullanımı-Mİ ilişkisiyle ilgili karar açısından çok anlamlı farklılıklara işaret etmediğini görüyoruz. Öte yandan, NSAİ ilaçlar jeneriğe göre incelendiğinde, gruplar arasında fark, hem vaka kontrol hem de kohort metaanalizlerinde dikkat çekiyor: Her iki çalışma grubunda da rofekoksib kullanımı, Mİ açısından daha riskli bulunmuş. Benzer bir bulgu, kolon adenomu hastalarında yapılan randomize kontrollü klinik çalışmalarda da bulunmuş. Burada Mİ açısından risk epeyce büyük. Diyorlar ki, “bu grupta zaten rofekoksib ve selekoksib kullandık, onun da yarısında selekoksibin günlük dozu 800 mg’ın üzerinde kullanıldığı için, bu sonuç olasıdır”. Bu tür farklılıkları göstermek açısından, metaanalizlerde alt grup analizleri yapmakta yarar var. Ancak, dikkat edilmesi gereken özellikler arasında, yapılacak alt grup analizi için örnek hacminin yeterli olması ve bu tür bir analizi gerektirecek bir biyolojik açıdan anlamlı hipotez olması önemlidir. Ne kadar çok alt grup analizi yapılırsa, gruba özel bir fark bulunma olasılığı o kadar artacağı için, kaç tane alt grup analizi yapılacağına özenle karar verilmelidir. Eğer gruplar için hesaplanan etken sonuç ilişkisi benzerlik gösteriyorsa, analiz için seçilen özelliğin olası karıştırıcı etkileri bilahare incelenebilir.

Son olarak, bir metaanalize ait tablo ve grafiği doğru okuyabilmek önemli. Örneğin, bu grafikte NSAİ ilaç-Mİ ilişkisini çalışan toplam 14 vaka kontrol çalışmada 74 673 Mİ hastası ve 368 968 kontrol incelenmiş. Sabit etki modeli kullanıldığında çok az bir risk fazlalığı var iken ($OR=1.32$, % 95 $GA=1.29-1.35$), rastgele etkiler modelinde bu riskin istatistiksel açıdan anlamlı olmadığına ($OR=1.17$; % 95 $GA=0.99-1.37$) dikkat çekmek istiyorum. Demin değindiğimiz gibi, eğer heterojenite söz konusu olduğunda sabit etki modeli kullanılır ise, hesap edilecek değer çalışılan ilişkileri istatistiksel açıdan “daha

anlamlı” gösterme eğiliminde sonuç verecektir buradaki fark da bu nedenle. Vaka kontrol çalışmaları bir araya getiren bu metaanalizde anlamlı heterojenlik saptanmıştır ($Ki-kare = 622$, $s.d.=13$, $p<0.00001$; I^2 değeri % 97’lerde). Öte yandan, bu heterojenliğin nereden kaynaklandığını göstermek de önemli. Alt grup analizlerinde, çalışmanın yapıldığı ülke, çalışmada örnek hacmi ve NSAİ ilaç alma durumunun nasıl tespit edildiği (reçete veya telefon görüşmesi) özelliklerine göre NSAİ ilaç kullanım-Mİ ilişkisi farklı özellikte ve değerde bulunmuş. Tablo 2’ye bakarsak, çok katılımlı çalışmalarda (kontrol sayısı >10 000), NSAİ ilaç kullanımında Mİ tahmini riski (*odds*) kullanmayan kişilere göre 1.34 kat (% 95 $GA= 1.18-1.51$) fazla iken, küçük çalışmalarda bu değer 0.70 (% 95 $GA= 0.39-1.27$) olarak bulunmuştur. Ayrıca, alt gruplar için kontrol edildikten sonra dahi çalışmalar arasında heterojenlik devam etmektedir ki, bu nedenle, rastgele etki modeli kullanılsa dahi verilecek tek bir *odds* oranı değeri uygun olmayabilir. Çalışma sonuçlarını ayrı ayrı bildirmek burada daha uygun olabilir.

Konuyu bitirmeden, kanıt-hasta-hekim üçlünün önemini vurgulamak isterim. Kanıt, ne kadar güçlü olursa olsun tek başına tedavi seçiminde karar verdirici olamayacaktır. Hastanın da konuyla ilgili bilgilendirilerek, karara ortak olmasının sağlandığı bir ortamda, ilaç seçiminin “hekim” tarafından “hasta” için, kanıtlar ışığında en uygun tarzda seçilmesi durumunda, ilacın etkililiğinin en yüksek olması beklenir. Dolayısıyla, kaliteli bir kanıt kadar, belki daha çok, kanıtın doğru ve uygun kullanılması da önem taşıyor. Burada biz, bu makaleyi tartışırken, kanıt-hasta-hekim üçgeninin sadece bir ayağı olan kanıtın kalite ve anlamından bahsedebiliyoruz.

Hasan Yazıcı: Başlangıçta heterojen bir grup biraraya gelip metaanaliz yapıyorlar. İki tane gözlemcileri var. Ve o gözlemcilerin aralarında anlaşmazlık olduğu zaman bir araya geldik diyorlar.

Banu Çakır: Buna ben de baktım, çünkü önemli bir konu. Yüzlerce çalışma incelemişler. Bu gözlemciler birbirinden bağımsız olarak, inceledikleri çalışmanın yapıldığı yıl, nerede yapıldığı, çalışmanın tasarımı, örnek büyüklüğü,

vaka kontrollerin nasıl seçildiği vb. bilgilerini inceliyorlar. Gözlemcilerin sadece iki çalışmada fikir ayrılıkları olduğu ve bir araya gelerek tekrar değerlendirme yaptıkları belirtiliyor sayfa 4'ün başında. Bu açıdan gözlemciler arası tutarlılık açısından çok huzursuzluk hissetmedim.

Bir dinleyici: Bir de vaka kontrol çalışmalara ait bir şekil var.

Banu Çakır: Bu grafiksel gösterim metaanalizlerde ağırlıklı olarak basıma bağlı (publication bias) ama diğer tipte de olabilecek bazı hata kaynaklarını tespit için yararlı. “*Funnel plot*”, metaanalize alınan çalışmaların her birinden elde edilen etken-sonuç ilişkisine ait etki gücünün, bu örnekte NSAI ilaç kullanımı-MI arasındaki ilişkiye ait *odds* oranı değerinin saçılım grafiği. X ekseninde etki gücüne ait değer verilmiş; burada *odds* oranı, ama farklı zamanlarda rölatif risk ya da OR veya RR'e ait logaritmik değer de kullanılabilir. Özet *odds* oranı değerinden x eksenine dikey bir çizgi çizilmiş. Y ekseninde çalışmaların örnek hacmi (sayı) görmeye alışkın, burada standart hata değeri verilmiş *odds* oranına ait standart hata, örnek hacmi büyüdükçe azalacaktır, dolayısıyla örnek hacmi yerine kullanılmış. Çalışmanın örnek hacmi büyüdükçe, çalışmadan elde edilen değer için özet değere yaklaşmasını bekleriz. Dolayısıyla, taraf tutma yok ise, küçük gruplarda yapılmış çalışmalardan elde edilmiş değerler, tepe noktası özet *odds* oranı değerinden başlayan ters bir V harfi şeklinde, grafiğin alt kısmında yaygın olarak saçılmış olacak; örnek hacmi arttıkça, çalışmaya ait *odds* oranı değerleri özet *odds* oranı değerine yaklaşmış görünümde olacaktır: bu görüntü bir külaha benzetilerek isimlendirilmiş bu grafik zaten. “*Publication bias*” kaynaklı olarak, eğer küçük sayıda kişi inceleyen ve negatif etki gösteren çalışmalar olması kadar yayımlanmamış ise, grafiğin sol alt kısmındaki noktalarda azalma bekleriz. Aslında, küçük örnekler ile ilgili çalışmalar ile ilgili diğer taraf tutmalar da bu tür grafikler yardımı ile incelenebilir. Burada sunulan grafik, NSAI ilaç-MI ilişkisi konusunda yapılmış küçük çaplı ve negatif etki saptanmış çalışmalar yönünde bir publication bias olduğunu düşündürüyor.

Sebahattin Yurdakul: Benzer konularda çifte metaanalizler oluyor ama hepsi farklı çıkıyor.

Yani ben kendimce benzer çalışmalara izah bulamıyorum. Acaba daha zekice bir izah tarzı var mı?

Banu Çakır: Bunun birçok nedeni olabilir. Seçtiğiniz çalışmaların özelliği (tasarımı, örnek büyüklüğü), çalışma grubunu nereden seçtiğiniz, o çalışmaları neden seçtiğiniz, yayımlanmamış çalışmaları dahil edip etmemiş olmanız, belirli bir dil dışı yayınları derlemeye dahil etmemek, metaanalize dahil etmeme kriterleriniz ve benzeri özelliklerin hepsi metaanalizden alacağınız sonucu etkiliyor. Bilinçli olarak bazı çalışmaları bir araya getirmenin yaratacağı hataları bir kenara bırakıyorum.

Bir de metaanalizin amacı önemli, örneğin, COX2-spesifik ilaçların sadece MI ile olan ilişkisi çalışılırsa, ilacın belirli bir hasta grubunda kullanılması hakkında negatif bir karar verebilirken; daha farklı bir metaanalizde COX2-spesifik ajanların gastrointestinal advers etkiler ile ilişkisi çalışıldığında farklı bir karar verilecektir. Bu anlamda, metaanalizlerde ilacın istenen ve istenmeyen ve farklı etkilerinin çalışılması önem taşımaktadır. Keza, ilaca ait maliyet-etkililik analizleri için de bu tür çalışmalara gereksinim gün geçtikçe artıyor. Geçen seneki metaanaliz konuşmasında Dr. İoannidis'in belirttiği bir konuyu burada vurgulamak yerinde olur. “Birileri bir şeyler yapsın ve ben bunu bekleyeyim de, metaanalizini yapayım, örnek hacmini büyüteyim” demek yerine gruplar, özellikle uluslararası işbirlikleri dahilinde bir araya gelip, herkes çalışmanın bir kısmını üstlenebilir ve tamamlandığında birleştirebilir. Bu tür ortaklıklarda seçilen hasta popülasyonlarının benzer olması, kullanılan yöntemlerin standart olması gibi özellikler de karşılanabilir ki, bu durumda, metaanalizlerin veri kalitesi de artacaktır.

Sebahattin Yurdakul: Peki, metaanalizlerin metaanalizi var mı?

Banu Çakır: Kümülatif metaanaliz buna benzer bir çaba. Kapsamlı bir metaanaliz sonrasında, zaman içinde yeni yayımlanan ya da yapılan çalışma sonuçları metaanalize eklenebilir. Cochrane grubunun bu tipte çalışmaları var biliyorsunuz; hazırlanmış metaanalizler düzenli olarak güncelleniyor,

yeni elde edilen kanıtlar mevcut metaanalize eklenebiliyor.

Bir dinleyici: Metaanaliz sonuçlarının çalışma grubuna benzemeyen toplumlarda ve farklı genetik yapısı olan gruplarda uygulanması bir hata değil mi?

Banu Çakır: Evet, ancak metaanalizin öyle bir iddiası yok zaten. Metaanalizde elde edilen veriler, olabildiğince tarafsız bir şekilde bir araya getiriliyor. Yani, okuyucu, birden çok makaleyi tek tek okuyup, sonuçlarını kafasında toparlamaya çalışacağına, hepsinin bir arada olduğu metaanalizi okuyor: çalışma gruplarının özellikleri, genellenebilirliği ile ilgili kısıtlılıklar metaanalizde irdeleniyor. Farmakogenetik ilacın etki ve güvenirliliğini çalışmada çok önemli. İlaç etkileri ve güvenliliği konusundaki Faz IV çalışmalar, gözlemsel araştırmalar, spontan advers etki bildirimleri bu nedenle çok değerli. Yine aynı nedenle, ilaç kullanımına ilişkin ulusal veri tabanlarının geliştirilmesi de çok yararlı. Farklı etnik, irksal grupları içeren metaanalizlerde alt grup analizi yapılabilir ama daha da doğrusu, böyle bir olasılık varsa, ilgilenilen ilişkileri özelliği olan gruplarda, ilişkilerin, uygun bir plan ile incelenmesi.

Yaşar Karaaslan: Metaanalizi yapan kişi, çalışmaları seçerken bazen negatif çalışmaları almayıp, pozitif çalışmaları da tekrar tekrar aynı çalışma farklı şekillerde üstüne biraz daha ilave olup yeni bir çalışmamış gibi alabiliyor; yani, sonuçta metaanalizleri yapan kişilerin dürüstlüğü denetleyen bir mekanizma var mı? Çünkü metaanalizi için 2 negatif çalışmayı almayıp 3 tane pozitif çalışmayı üç kere koyarsanız, istediğiniz sonuç çıkar.

Vedat Hamuryudan: QUOROM bildirisi var. Önemli dergilerde bir metaanaliz yayımlanacağı zaman bu bildirideki kurallara uymak gerekiyor. Bu da bir kalite kontrolü getiriyor. Hangi çalışmaları aldın? Hangilerini almadın? Sadece senin işine gelenleri mi aldın gibi açıklanması gereken sorulardan oluşuyor.

Banu Çakır: Farklı araştırmaların ve yazıların okunup incelenmesi, derlenmesi konusunda kullanılabilecek bazı rehberler var. Randomize kontrollü çalışmalara ait metaanalizler için QUOROM, gözlemsel araştırmalara ait

metaanalizler için MOOSE, tanısal çalışmaların değerlendirilmesinde STARD gibi.

Yaşar Karaaslan: Biri yazı gönderdi, bu yazıda serotonin geri alım inhibitörleri ile ilgili yapılmış bir metaanalizin dökümü yapıyordu. Negatif çalışmaları almayın, pozitif çalışmalar endüstri destekli; farklı farklı uzatmış. Biraz da hasta koymuş, ikisini birleştirmiş ve onları 5 ayrı çalışma gibi koyup, negatif çalışmaları da bulanabileceği halde metaanalizi almamışlar. Metaanalizi eleştiren bir yazıyı Vedat hocam bana göndermişti. Eskiden çalışma okurken, *conclusion* kısmını okuyup çalışmayı okudum zannedirdim. Şimdi metaanalizlerin bana elma ve armudun toplanması gibi geliyor. Yani ben metaanalize hiç kıymet vermiyorum açıkçası, bilmiyorum yanlış mı yapıyorum.

Banu Çakır: Kötü örnekler olabilir. Ama iyi yapılırsa, metaanalizler kıymetli. Tek bir çalışmada bile gerçeği bulmayı engelleyen çok sayıda bias ve karıştırıcı faktör rol oynayabilir. Kaldı ki, her birinin kısıtlılığı olan bir grup çalışmayı bir araya getirmek bazen mümkün değil, mümkün olduğunda da mutlaka uzman bir ekip tarafından yapılması uygun. Bu ekipte, ilgili klinisyen, epidemiyolog, istatistikçi ve diğer ilgili kişiler birlikte çalışmalı. Unutulmaması gereken konu, kumaş ne kadar iyiye dikilecek kıyafetin o kadar iyi olabileceği.

Söylediğiniz diğer konu çok önemli. Çalışmaların önce yöntem kısımları okunmalı. Ancak yöntemi okuyan bir kişi çalışmanın devamını okuyup okumama kararını verebilir ya da okuduğu "kanıt" ne kadar geçerli olduğuna veya bu bilgiyi kendi hastasına uygulayıp uygulayamayacağına karar verebilir.

Tuulikki Sokka: Dün bahsettiğim çalışma için Finlandiya popülasyonu veritabanlarını kullandıklarını biliyorum ve özellikle Finlandiya'da bu sonuçların doğru olduğuna inanıyorum. İskandinav'da ise metaanalizin toplam sonuçlarından çok onun sonuçlarına güvenirliler. Bu belki sizin ilginizi çekmez ama durum böyledir.

Banu Çakır: Eğer kendi ülkenizde yapılmış yöntemsel açıdan güvenilir ve yeter büyüklükte bir çalışma varsa, kendi genetik yapınızı yansıtaacağı ve ilaç kullanımını etkileyen bir çok

sosyal ve kültürel özellik açısından hasta popülasyonunuza yakın olacağı için bu çalışma sonuçlarına bakmakta yarar var. Ama ülkesinde çalışma yapamayanlar nereye bakacaklar? Kaldı ki, farklı toplumlarda yapılan çalışma sonuçları, ilaç ile ilgili bilgilerimizi her koşulda zenginleştirecektir. Bu anlamda da, metaanalizler önemli.

Bir dinleyici: Kanımca metaanalizleri klinik çalışmaların yeterli sayıya ulaşmaması nedeni ile güç artırarak anlamlılık aramak şeklinde algılasak, bu son derece yanıltıcı ve zararlı bir yaklaşım olur. Bence metaanalizleri bir alandaki kanıtları olabildiğince objektif ve saydam şekilde enine boyuna kavramak şeklinde algılamalıyız. Daha ideal olanı alanın sınırlarını bilmeliyiz. İnsanların kendi çalışmalarını yapmasını beklemek, ardından bilgileri olaydan sonra toplamak yerine, herkesin ilgisini çeken sorular üzerinde dünya çapında araştırmacı ekipler kurup, prospektif metaanaliz 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 yayımlayabilir ve buluşun övgüsünü toplayabilir. Ancak, kalan 99 ekibin bilgisi ile biz bu bilgiyi sistematik olarak tekrarlayıp geçerliliğini sıkı metaanalizleri yöntemleriyle sınavabileceğimizi de biliriz. Yani metaanaliz kavramı yavaş yavaş uluslararası işbirliği ve prospektif tasarımlara doğru kaymaktadır.

Kaynaklar:

1. Downs SH, Black N. The feasibility of creating a check list for the assessment of the methodological quality both of randomized and nonrandomized studies of health care interventions. *J Epidemiol Community Health* 1998; 52: 377 – 384.
2. Clark HD, Wells GA, Huet C ve ark. Assessing the quality of randomized trials: reliability of the Jadad scale. *Control Clin Trials*. 1999; 20: 448 – 452.
3. Matsuda T, Ohno S, Hirohata S ve ark. Efficacy of rebamipide as adjunctive therapy in the treatment of recurrent oral aphthous ulcers in patients with Behçet's disease: a randomised, double-blind, placebo controlled study. *Drugs R D*. 2003; 4: 19 – 28.
4. Masuda K, Nakajima A, Urayama A ve ark. Double masked trial of cyclosporin versus colchicine and longterm open study of cyclosporin in Behçet's disease. *Lancet*. 1989; 1: 1093 – 6.
5. Scott PA, Kingsley GH, Smith CM, Choy EH, Scott DL. "Nonsteroidal antiinflammatory drugs and myocardial infarctions: comparative systematic review of evidence from observational studies and randomised controlled trials" *Ann Rheum Dis*. 2007;66(10):1296304.

Daha fazla bilgi için önerilen kaynaklar:

1. Strom BL, Kimmel SE (eds). *Textbook of pharmacoepidemiology*. John Wiley & Sons, Ltd., NJ, U.S.A., 2006.
2. Elwood M. *Critical appraisal of epidemiologic studies and clinical trials*. Third edition. Oxford, New York, NY, 2007.
3. Pocock SJ. *Clinical trials. A practical approach*. John Wiley & Sons, Ltd., New York, NY, 1983.
4. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of clinical trials*. Third edition. Mosby Year Book, Inc., 1996.
5. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. "Improving the quality of reports of meta analyses of randomized controlled trials: The QUOROM Statement. Quality of Reporting of Meta Analyses. *Lancet* 1999;354:18961900. (Not: yakın zamanda PRISMA olarak isimlendirilmiştir, bu isim ile de araştırabilirsiniz).
6. Stroup DF, Berlin JA, Morton SC, Ingram O, et al. "Metaanalysis of observational studies. A proposal for reporting". *JAMA*, 2000;283(15):200812. (Not: kısa adı MOOSE)
7. Bossuyt PM, Reitsma JB, Bruns DE, et al. . "Towards Complete and Accurate Reporting of Studies of Diagnostic Accuracy: The STARD Initiative". *Clinical Chemistry* 2003;49 (1):1–6

A

- ACR, bkz. American College of Rheumatology
 AIMS, bkz. Artrit etki ölçümü skalaları
 American College of Rheumatology 27, 28,
 29, 30, 31, 32, 46, 70, 74, 78, 97
 anti-CCP 29, 40, 48
 ARA, bkz. Amerikan Rheumatism Association
 ARMD veri tabanı 75
 artrit etki ölçümü skalaları 30

B

- bias 36, 97, 101, 102
 biyomedikal model 29
 BOARD veri tabanı 75

C

- causal, bkz. nedensel
 CDAI, bkz. klinik hastalık aktivitesi
 Cochrane 98, 101
 cohort 98
 compliance 98
 CONSORT 98

D

- DAS, bkz. hastalık aktivite skoru
 daha az etkili olmama 9, 12, 15, 16, 17, 18, 19,
 20, 21

E

- EMEA 11, 19
 equivalence 9, 15
 EuroQol, bkz. Avrupa yaşam kalitesi indeksi
 extension studies, bkz. uzatma çalışmaları

F

- FDA 11, 19, 24, 37, 73, 97
 fonksiyonel kapasite 35, 60, 61
 fonksiyonel maluliyet 30, 38

G

- genetik analiz programı 87
 genetik ilişki çalışmaları 81, 82, 86, 87, 91, 92
 genetik model 87, 89
 genetik polimorfizm 89
 geri ödeme 59, 61
 global değerlendirme 18, 76, 95, 96

- global bir veri tabanı 67
 gözlemsel araştırmalar 98, 102
 güç 12, 25, 34, 41, 81, 82,
 83, 84, 85, 86, 87, 88, 89, 91, 94, 97, 98, 103
 güven aralığı 13, 17, 64, 65, 91, 95, 99
 güvenlik intervalleri 48

H

- hastalık aktivite skoru 27, 28, 30, 31, 32, 39, 40,
 69, 40, 69, 70, 74, 76, 77, 78
 hastanın kendisinin bildirdiği 28, 32, 38
 HAQ, bkz. RA
 head-to-head bkz. kafa kafaya
 Helsinki Deklarasyonu 96
 HWE, bkz. hardy-weinberg dengesi

I

- Informed consent 97
 Intention to treat 20, 25, 95
 ITT 20, 21

K

- kantitatif hasta anketleri 31, 33
 kesme noktası 19, 20
 ki kare testi 10, 12, 13, 16, 96
 kohort, bkz. cohort

- kohort çalışmaları 97, 98
 Kuruskal-Wallis testi 10
 kümülatif metaanaliz 98, 101

L

- Le Henanff 21

M

- MACTAR, bkz. McMaster Toronto artrit hasta
 tercihi anketi
 Mc Nemar testi, bkz. ki kare testi
 McMaster Toronto artrit hasta tercihi anketi 30
 MDHAQ 30, 31, 32, 37, 38, 39, 41,
 73, 74, 76, 77, 79
 metaanaliz 13, 18, 20, 34, 91, 93, 97,
 98, 99, 100, 101, 102, 103
 MI 11, 12, 16, 18, 19
 99, 100, 101

mortalite 11, 12, 16, 17, 18, 19, 35, 60, 61, 74, 77

N

Nedensel 24, 25, 98

neoplazi 7, 25, 64, 65

non- inferiority, bkz. daha az etkili olmama

NSAID 19, 34

non-hypothesis, bkz. yokluk hipotezi

null-hypothesis 10, 11, 12, 15

O

odds ratio, bkz. tahmini rölatif risk

O'Brien- Fleming 12

P

PAWE, bkz. power for association with error

PET, bkz. hasta değerlendirme tekniği

P değeri 9, 10, 11, 13

PL, bkz. plasebo

plasebo 9, 12, 13, 15, 19, 20, 23, 24, 28, 32, 33,

35, 36, 37, 40, 41, 82, 83, 93, 94, 96, 97

polimorfizm 88, 89, 90, 91, 92

power, bkz. güç

power and sample size calculation 85

power for association with error 86, 87

prevalans 30, 61, 95

PP, bkz. per-protocole

publication bias, bkz. yayın saptırması

Q

Quest – RA, bkz. global bir veri tabanı

QUOROM 102

R

RA 7, 8, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 40, 60, 61, 62, 67, 68, 69, 70, 71, 73, 74, 75, 77, 78

RADAI, bkz. hastanın kendisinin bildirdiği

romatoid artrit hastalık aktivite indeksi

RADAR, bkz. romatolojide hastalık aktivitesinin

hızlı değerlendirilmesi

randomize kontrollü çalışma 10, 13, 15, 16, 19,

21, 23, 32, 33, 34, 35, 36, 41, 82, 97, 100, 102

RAPID 30, 32, 37, 38, 39, 40, 41, 43, 47, 73, 74, 76, 78

rastgele etkiler modeli 99, 101

RCT (RKÇ), bkz. randomize kontrollü çalışma

relative risk scale 82

rescue 96

risk 11, 19, 20, 24, 37, 63, 64, 65,

82, 83, 85, 91, 98, 99, 100, 101

Ritchie Endeksi 27

rölatif risk 82, 83, 98, 101

run simulation 83, 84

S

Sharp skoru 28

SIR değeri 64, 65

sistemik derleme 97, 99

sistemik hata 98, 99

SNP 89, 90, 91

standart hata 95, 98, 101

standart sapma

STROBE 98

superiority, bkz. üstünlük

surrogate marker, bkz. vekil belirteç

T

temel veri seti 27, 28, 30, 32, 35, 39

T-testi 11, 16

tip I hata 11, 12, 99

tip II hata 11, 12

TNFRSF 91

U

uzatma çalışmaları 24, 25

Ü

üstünlük 9, 12, 15, 16, 17, 18, 20, 21, 28, 83

V

vaka kontrol 86, 97, 98, 99, 100, 101

varyans analiz testleri 10

VAS 9, 30, 31, 32, 38, 39, 41, 74

vekil belirteç 35

veri tabanı 7, 27, 33, 48, 59, 60, 61, 62,

67, 70, 71, 73, 77, 94, 97, 102

Y

yanlış negatif	29
yanlış pozitif	11, 29
yokluk hipotezi	10, 11, 12, 15
yuvalandırılmış vaka kontrolü	98

W

Western Ontario McMaster osteoartrit indeksi	30
WOMAC, bkz. Western Ontario McMaster osteoartrit indeksi	
WOMACPA	9, 11, 13, 18
WOMACPH	9, 11, 18
Wilcoxon testi	10