

# ARAŞTIRMA MAKALELERİNİN KRİTİK ANALİZİ

Barış ATA, Bülent URMAN

V. k. v. Amerikan Hastanesi, Kadın Sağlığı Merkezi, Tüp Bebek Ünitesi

## ÖZET

*Teknoloji ve iletişimdeki ilerlemeler sonucunda giderek artan hız ve sayıda araştırma makalesi yayınlanmaktadır. Maalesef en yüksek standartları korumaya çalışan uluslararası hakemli dergiler de dahil olmak üzere yayınlanan araştırmaların çoğunun metodolojik kalitesi yüksek değildir. Ülkemizdeki tıp eğitimi sürecinde metodolojiye ayrılan zamanın azlığı da bu faktörlere eklendiğinde uygulamalarını kanıta dayalı tıp prensipleri ışığında yönlendirmek durumunda olan doktorların karar verme sürecinde kendilerine yardımcı olacak gerçek ve uygun kanıtları bulup değerlendirmesi zorlaşmaktadır. Bu makalede yayınlanmış araştırma makalelerinin kritik analizinde göz önüne alınması gereken metodolojik kavramlar özetlenmiştir.*

**Anahtar kelimeler:** istatistik yöntemler, klinik araştırmalar, kritik analiz, metodoloji

*Türk Jinekoloji ve Obstetrik Derneği Dergisi, (TJOD Derg), 2008; Cilt: 5 Sayı: 2 Sayfa: 83- 8*

## SUMMARY

### Critical appraisal of trial reports

*An increasing number of research articles are being published every day due to rapid developments in technology and communication. Unfortunately, most of these published articles are far from reaching the highest methodological quality, even in most prestigious international peer reviewed journals. Considering the limited time devoted to methodological training during medical education on top of these facts, sifting the correct and relevant evidence from the pile of published material is a difficult task for the clinician who is expected to practice in the light of evidence based medicine. Methodological issues in critical appraisal of published research articles are summarized in this review.*

**Key words:** clinical trials, critical appraisal, methodology, statistical methods

*Journal of Turkish Obstetric and Gynecology Society, (J Turk Obstet Gynecol Soc), 2008; Vol: 5 Issue: 2 Pages: 83- 8*

## GİRİŞ

“Kanıt dayalı tıp” (KDT) kavramı 1990lı yıllardan itibaren tıbbi uygulamaları yönlendirmektedir. Tıbbın tüm alanları gibi obstetrik ve jinekoloji alanındaki uygulamalar da KDT prensipleri dahilinde gerçekleştirilmelidir. Maalesef en yüksek standartlardaki dergilerdekiler de dahil olmak üzere yayınlanan çalışmaların çoğu metodolojik açıdan yüksek kalitede değildir<sup>(1)</sup>. Gerek giderek artan hız ve çoklukta yayın-

lanan verilerin varlığı, gerek yayınlanan çalışmaların düşük kalitesi, gerek ülkemizdeki tıp eğitimi sürecinde verilen metodolojik eğitimin azlığı diğer alanlarda çalışan meslektaşlarımız gibi biz jinekologların da uygulamalarımızı dayandıracığımız, karar verme aşamalarında bize yardımcı olacak gerçek ve uygun kanıtları bulmamızı ve değerlendirmemizi zorlaştırmaktadır. Bu makalede yayınlanan çalışmaların metodolojik açıdan sorgulayarak eleştirel gözle okunmalarına yardımcı olacak temel kavramlar özetlenmiştir.

**Yazışma adresi:** Dr. Barış Ata, Güzelbahçe Sokak No: 20 Nişantaşı, 34365 İstanbul

Tel.: (0532) 744 74 16

e-posta: barisata@hotmail.com

Alındığı tarih: 09.04.2008, revizyon sonrası alınma: 09.04.2008, kabul tarihi: 02.05.2008

Bir araştırma makalesini okurken amacımız söz konusu çalışmanın bulgularına göre kendi uygulamamızı yönlendirmektir. Yayınlanan çalışmaların çokluğu göz önüne alındığında tümünü okumamızın ve sonradan içeriğinin bizim için gerekli ve uygun olup olmadığını değerlendirmemizin mümkün olmadığı açıktır. Bu nedenle dikkatimizi çeken çalışmaların özetini okuduktan sonra tam metnini okumaya zaman ayırıp ayırmayacağımıza karar veririz. Bir araştırma makalesini değerlendirirken yayınlandığı dergi, yazarlarının akademik bilinirlik ve güvenilirliği de şüphesiz önemlidir ancak burada sadece metodolojik konular ele alınacaktır.

### **Çalışmanın amacı**

*Amaçta dört noktanın açıkça belirtilmiş olması gereklidir; i) incelenen popülasyon, ii) karşılaştırılan tedavi veya girişimler, iii) araştırmanın tasarımı ve iv) sonuç ölçütü olarak alınan klinik durum. Bu dört noktanın her birisi makalede verilen sonuçların bizim kendi uygulamamız için bizim çalışma koşullarımızda ne kadar tekrarlanabilir ve geçerli olabileceğini değerlendirmekte önemlidir.*

Kaliteli bir araştırmanın amacı çok uzun olmayan bir cümle ile açıkça ifade edilebiliyor olmalıdır. Ancak bu şekilde araştırmanın sonuçları karışıklığa neden olmadan anlaşılabilir.

İncelenen popülasyonun demografik, sosyal veya biyolojik özellikleri, deneklerin hastalıklarının durumu bizim hizmet verdiğimiz popülasyondan önemli ölçüde farklılık gösteriyorsa araştırmanın sonuçları bizim koşullarımızda geçerli olmayabilir.

Çalışma grubunda uygulanan tedavi veya girişim gerek tedaviye erişim, gerek ekonomik koşullar, gerek hastalarımızın ve bölgemizin sosyal koşulları nedeniyle bizim uygulayabildiğimiz veya uygulayabileceğimiz bir yöntem değilse çalışmanın sonuçları yine en azından günlük uygulamamızda işimize yaramayabilir. Kontrol grubunda uygulanan ve aktif tedavinin karşılaştırıldığı tedavi veya girişimin doğru seçilmiş olması çok önemlidir. Kontrol grubunda uygulanan tedavi incelenen hastalığın tedavisinde güncel koşullarda yaygın kabul gören ve etkinliği kanıtlanmış bir yöntem olmalıdır ki deneysel tedavi veya girişimin mevcut standarda göre üstünlüğü, kolaylığı, yan etkileri ve/veya maliyeti doğru olarak karşılaştırılabilir olsun. Kontrol grubunda uygulanan tedavi veya girişim hastalığın doğal seyrinde anlamlı bir değişiklik yaratmıyorsa çalışma grubunda uygulanan tedavinin

etkisi gerçekte olduğundan daha fazla görünecektir. Sonuç ölçütü olarak değerlendirilen kriter net olarak belirlenmiş olmalıdır. Örneğin, infertilite çalışmalarında değerlendirilmesi gereken sonuç ölçütü neredeyse her zaman en azından devam eden gebelik olmalıdır<sup>(2)</sup>. Gelişen folikül sayısı, elde edilen oosit sayısı, sperm parametrelerindeki değişiklikler, embryo implantasyon oranları gibi ölçütler bazı koşullarda akademik anlamda fikir verici olabilirler de tedavi etkinliği açısından önemli olan sonuç ölçütü gebeliktir. Yukarıda örnekleri verilen parametrelerdeki değişiklikler her zaman gebelik oranı veya gebelik elde edene kadar geçen zaman gibi klinik önem taşıyan parametrelere aynı büyüklükte yansımayaabilmektedir<sup>(3)</sup>. Tedavilerin etkinlikleri dışında yine klinik uygulamada önem taşıyan istenmeyen etkileri, yan etkileri, komplikasyon oranları gibi parametreler de değerlendirilmeli ancak bu sonuçlar yorumlanırken doğru yöntemler titizlikle seçilmeli ve yorumlara dikkat edilmelidir. Bir çalışmada elde edilen tüm bulgular değerlendirilmeye ve her bulgu için olası nedenler açıklanmaya çalışıldığında genellikle çalışmanın asıl amacından uzaklaşarak aşağıda örneklem hacmi konusunda tartışılacak olan istatistik nedenlerden dolayı yetersiz ve yersiz çıkarımlar yapılmış olması olasılığı artacaktır.

## **GEREÇ ve YÖNTEM**

### **Randomizasyon**

*Randomizasyon deneklerin çalışmaya katıldıktan sonra hangi gruba ayrılacaklarının rastlantısal olarak belirlenmesi demektir. Deneklerin çalışma ve kontrol gruplarına rastgele ayrılması çalışma gruplarının etkinliği incelenecek olan tedavi dışındaki öngörülen ve öngörülmeleyen tüm faktörler açısından benzer özelliklere sahip olmasını sağlamak amacıyla yapılır.* Değerlendirilen sonuç parametresinde gözlenen bir farkın başka bir faktöre değil de incelenen tedaviye bağlı olduğu sonucuna ancak bu şekilde varılabilir. Randomizasyon gerek bazal parametrelerde gerekse sonuç ölçütünde tesadüfi bir farkın asla olmayacağını garanti edemez. Ancak araştırmacıların bilinçli veya bilinçsiz olarak daha iyi prognozlu hastaları bir gruba kötü prognozlu hastaları diğer gruba yönlendirmesine engel olur.

Randomizasyon çeşitli yöntemlerle yapılabilir. Basit randomizasyon her yeni hasta çalışmaya katıldığında

hangi gruba ayrılacağına yazı tura atarak karar verilmesidir. Basit randomizasyon uygulandığında çalışma tamamlandığında iki gruptaki deneklerin sayıları arasında büyük farklılık olabilir. Çalışma ve kontrol gruplarındaki hasta sayıları arasındaki fark büyüdükçe yapılacak istatistik analizin güvenilirliği azalacaktır. Bu durumu önlemek için kullanılacak iyi bir yöntem blok randomizasyondur<sup>(4)</sup>. Blok randomizasyonda tüm deneklerin sayısı bloklara bölünür ve her blok içerisindeki hastalar çalışma ve kontrol gruplarına eşit olarak dağılır. Sonuçta çalışma tamamlandığında her iki gruptaki hasta sayısı yaklaşık aynı olacaktır.

İdeal yöntem katılan heryeni hastanın hangi grupta olacağına hasta kaydı başlamadan önce hazırlanan blok randomizasyon listesi araştırmacı ve hastalardan gizli tutularak üçüncü bir şahıs tarafından bildirilmesidir. Bu yöntemin bir faydası da çalışmaya bir kez katılan bir hastanın grubunun ancak çalışmaya kaydı yapıldıktan sonra belirlenmesi ve araştırmacı veya hastanın çalışma grubuna göre katılımdan vaz geçmesinin önlenmesidir. Hastaların dosya numaraları, isimlerinin harfleri, baş vuru günleri gibi öngörülebilir belirteçlere göre gruplara ayrılması yöntemlerine genel olarak quasi-randomisation (randomizasyon benzeri yöntem) denir. Bu yöntemler uygun koşullarda doğru uygulandıklarında oldukça basit ve etkin olmalarına rağmen prensip olarak çalışmaya katılması olası bir denneğin katıldığı takdirde hangi gruba gireceği araştırmacı tarafından öngörülebileceği için bilinçli veya bilinçsiz olarak yönlendirmeye açıktır. Gerçek randomizasyon yapılmayan çalışmalar, veya randomizasyon yöntemi açıkça belirtilmeyen çalışmalar yönlendirmeye açık olabilir ve sonuçları şüpheyle karşılanmalıdır.

### **Randomizasyon bilgisinin gizlenmesi (Allocation Concealment)**

*Potansiyel deneklerin çalışmaya katılmaları halinde hangi grupta yer alacakları veya çalışma sırasında hangi grupta yer aldıkları bilinirse sonuçlar bilinçli veya bilinçsiz olarak yönlendirilebilir. Bu koşul hem denegi çalışmaya kaydeden araştırmacı, hem tedaviyi uygulayan araştırmacı, hem denegin kendisi hem de sonucu değerlendiren araştırmacı için geçerlidir. Çalışmaya katılma kriterlerine uyan bir hastayla görüşen araştırmacı eğer denegin hangi çalışma grubuna gireceğini önceden bilir veya kestirebilirse kendi fikrine göre o hastayı çalışmaya hiç katmamayı seçebilir. Örneğin*

araştırılan yeni bir ilacın kontrol grubundaki tedaviye kıyasla daha iyi olacağını düşünen bir araştırmacı sempati duyduğu bir hastayı kontrol grubuna gireceğini bildiği takdirde çalışmaya dahil etmeyip ona deney grubundaki tedaviyi uygulamayı tercih edebilir. Alternatif bir senaryo ise daha iyi prognozlu hastalar bir gruba kötü prognozlu diğer gruba denk gelecek şekilde çalışmaya dahil edilebilirler. Buna seçim yanlılığı (selection bias) denir. Seçim yanlılığını önlemenin yollarının başında çalışma süreci boyunca katılım kriterlerine uyan her hastanın ardışık olarak çalışmaya dahil edilmesi gelir. Bu uygulama makalede açıkça belirtilmiş olmalıdır. Randomizasyon listesi araştırmacı ve hastalardan gizlenerek her bir denegin dahil olacağı grubun ancak o denek çalışmaya dahil olduktan sonra çalışmada direkt bir rolü olmayan üçüncü bir şahıs tarafından araştırmacıya bildirilmiş olması ideal yöntemdir. Grup dağılımının önceden hazırlanan numaralı opak zarflar içerisinde araştırmacıya önceden verilmiş olması açık nedenlerden dolayı ideal bir yöntem değildir. Araştırmacı önce zarfı açtıktan sonra hastayı çalışmaya kaydedip etmemeye karar vermiş olabilir.

Bir araştırma makalesinde randomizasyon yönteminin belirtilmesi kadar çalışma gruplarına dağılımın araştırmacılara nasıl bildirildiğinin belirtilmesi da önemlidir<sup>(4)</sup>.

### **Maskeleye (Körleme, blinding)**

*İdeal yöntem araştırmacıların ve deneklerin uygulanan tedavi protokolünü bilmediği çift kör tasarım olmasına rağmen bu uygulama her zaman mümkün olmayabilir. Bu durumda en azından sonuçları değerlendirenlerin çalışma grubunu bilmemesi sağlanmış olmalıdır.*

Bir hastanın uygulanan bir tedavi veya girişime yanıtını belirleyen tek faktör ilacın içeriği veya girişimin kendisi değildir. Mekanizması henüz tam olarak açıklanmış olmasa da “plasebo etkisi” mevcuttur<sup>(5)</sup>. Uygulanan tedavinin biyolojik bir etkinliği olmadığı halde kendisine fayda sağlayacağını düşünen hastalarda, özellikle subjektif değerlendirmelere dayanan sonuç ölçütlerinde iyileşmeler gözlenebilmektedir. Sonuçları etkileyebilecek bir diğer fenomen de “Hawthorne etkisi”dir<sup>(6)</sup>. Hawthorne etkisi terimi gözlem altında olduğunu bilen deneklerin davranışlarında oluşan değişiklikleri tanımlamak için kullanılır. Bunlara ek olarak bir denegin hangi grupta olduğunu bilmesi çalışma protokolüne uyuncunu etkileyebilir. Aktif tedavi grubuna katılacağı düşüncesiyle çalışmaya dahil olan bir denek plasebo grubunda olduğunu öğrenmesi halinde çalışmayı terk edebilir, aktif gruptaki

ilaç ulaşılabilir ise kendisi alıp kullanabilir. Araştırmacı açısından bakıldığında ise hastanın grubunu bilmesi halinde aktif tedavi grubundaki bir hastayla daha yakından ilgilenebilir, tedavisinde ve takibinde daha özen gösterebilir. Ölçümler sırasında rakamları farklı yönlerde yuvarlayabilir, veya subjektif değerlendirmelerde bir grup lehine davranabilir. Sonuç ölçütlerini değerlendirenler de benzer şekillerde davranabilir. Bu nedenle mümkün olduğu sürece deneklerin gruplara dağılımı randomizasyon kodu açıklanana kadar ne araştırmacılar ne de denekler tarafından anlaşılacak şekilde çalışma tasarlanmış olmalıdır. Bazı koşullarda maskeleyen sağlanması zor olabilir ve pratik nedenlerle tercih edilmemiş olabilir. Farklı özelliklere sahip ilaçlar karşılaştırılıyorsa double dummy tasarım uygundur. Yani oral yolla kullanılan bir ilaç ile injeksiyon şeklinde kullanılan bir ilaç karşılaştırılıyorsa ideal uygulama bir gruba aktif tablet ve plasebo injeksiyon diğer gruba plasebo tablet ve aktif injeksiyon verilmesidir. Günlük uygulama sayısı farklı ilaçlar kıyaslanıyorsa da aynı koşul geçerlidir. Günde bir kez alınan A ilacıyla günde üç kez alınan B ilacı karşılaştırılıyorsa, bir gruba günde bir kez A ilacı yanında günde üç kez plasebo tablet, diğer gruba günde bir kez plasebo yanında günde üç kez B ilacı verilmiş olmalıdır.

Cerrahi girişimleri değerlendiren araştırmalarda prensip farklı olmamakla beraber plasebo cerrahi (sham operasyonu) uygulamanın kendine has riskleri olduğundan etik koşullar dikkate alınmalıdır. Sham operasyon uygun görülmediğinde tercih edilecek bir yöntem aktif cerrahi uygulanan bir grup hastanın sonuçlarıyla daha önce alternatif cerrahi yöntemin uygulanmış olduğu hastaların sonuçlarını karşılaştırmak olabilir. Bu durumda seçilen kontrol grubunun çalışma grubuna benzerliği ve nasıl derlendiği dikkatle incelenmelidir.

### **Eşlik eden uygulamalar**

*Prensip olarak iki grup arasındaki tek fark sadece ve sadece araştırılan tedavi veya girişim olmalıdır.* Randomize bir çalışmanın amacı iki grup arasındaki tek farklılığın sadece etkinliği araştırılan tedavi olmasıdır. Bir gruba çalışmanın amacında belirtilen uygulamanın yanında ek bir ilaç daha uygulanmışsa sonuçta gözlenen farkın araştırılan tedaviye mi yoksa ek uygulama mı bağlı olduğu bilinemez. Örnek olarak iki farklı gonadotropin türünün sonuç üzerindeki etkisi

incelenirken bir gruba farklı bir luteal destek protokolü uygulanmışsa gebelik oranları arasında gözlenen farkın sadece gonadotropin türünden kaynaklandığından emin olunamaz.

Uygulama farklılıkları sadece verilen ilaçlar olarak algılanmamalıdır. Eğer bir gruptaki denekler diğer gruba kıyasla daha sık monitörize edilmişse, farklı bir ekip tarafından tedavi edilmişse de sonuçlar farklı olabilir.

Kaliteli bir makalede hem çalışma grubuna hem de kontrol grubuna uygulanan tüm tedavi ve izlem protokolleri açık bir şekilde, eksiksiz ve anlaşılabilir olarak belirtilmiş olmalıdır.

### **Örneklem hacmi hesaplaması, kullanılan istatistik yöntemler ve istatistik sonuçların yorumu**

*Küçük örneklem hacimli araştırmaların sonuçlarının gerçeği yansıtmaktan çok tesadüfi olma olasılığı çok yüksektir.*

İki grup arasında sonuç ölçütünde gözlenen fark ya gerçekten tedavinin etkinliğine ya da tesadüfe bağlıdır. Metodolojik açıdan kusursuz bir araştırmacının sonucunda da sadece tesadüfi farklılıklar gözlenebilir. Bir araştırma sonunda gözlenen boyutta veya daha yüksek bir farkın tesadüfen gerçekleşmiş olması olasılığı p değeri ile ifade edilir. Örneğin A ve B tedavileri sonunda A grubunda iyileşme oranı %20, B grubunda ise %27 bulunmuş ve  $p = 0.03$  olarak hesaplanmış olsun, bu sonuç A ve B tedavileri arasında iyileşme sağlamak açısından hiç bir fark olmasaydı, aynı sayıda hasta içeren bir araştırmada iki grup arasında sadece tesadüfen %7 (%27 - %20) veya daha yüksek bir fark gözleme olasılığı %3'tür anlamına gelir.

İki tedavi arasında gerçekte etkinlik farkı yokken çalışma sonucunda fark bulmak tip 1 hata veya  $\alpha$  hata olarak tanımlanır, ve p değeri tip 1 hata olasılığını belirtir. Genel olarak %5'in altında ( $<0.05$ ) p değeri istatistiksel olarak anlamlı kabul edilir. Araştırma sonuçlarının yorumunda p değeri kadar dikkat edilmesi gereken bir nokta iki grupta gözlenen mutlak değerlerdir. İstatistiksel anlamlılık her zaman klinik anlamlılıkla eş değildir. Örneğin bir çalışmada A ve B tedavileri sonucunda sırasıyla %30 ve %31 iyileşme oranı elde edilmiş olsun. Eğer çalışma yeterince çok sayıda hasta içeriyorsa gözlenen mutlak değer olarak %1 fark istatistiksel anlamlılığa ulaşabilir yani p değeri 0.05'ten küçük olabilir. Ancak B ilacı A ilacından daha fazla yan etki gösteriyorsa veya çok daha pahalıysa %1 daha fazla iyileşme sağlamak için klinik uygulamada

tercih edilmeyebilir. Bu durumda fark istatistiksel olarak anlamlı olsa bile klinik olarak anlamsız olacaktır. Farklı ve daha yaygın bir sorun, incelenen iki tedavinin etkinliği gerçekte farklıyken araştırma sonunda iki tedavinin etkinliği arasında fark gözlenmemesi veya gözlenen farkın istatistiksel olarak anlamlılığa ulaşmamasıdır. Bu durum tip 2 hata veya  $\beta$  hata olarak adlandırılır. Çalışmanın gücü  $1 - \beta$  hata olasılığı olarak ifade edilir ve genel olarak araştırmaların en az %80 güce sahip olması gerekir. Bunun anlamı gerçekte iki tedavi arasında fark olduğunda araştırma sonucunda bu farkı ortaya koyabilme olasılığının en az %80 olması demektir. Diğer bir deyişle iki tedavi arasında gerçekte var olabilecek bir farklılığı ortaya koyamama olasılığı %20'nin altında olmalıdır.

Randomize kontrollü etkinlik çalışmaları yapılırken kontrol grubunda beklenen pozitif sonuç oranı (event rate), yeni tedavi ile bu oranda elde edilmesi beklenen artış (klinik olarak anlamlı kabul edilecek en küçük oran alınmalıdır), tip 1 hata olasılığı (genellikle %5) ve tip 2 hata olasılığı (genellikle maksimum %20) göz önüne alınarak her grupta en az kaç hasta olması gerektiği hesaplanmış ve bu hesaplamada kullanılan rakamlar materyal ve metodda açıkça belirtilmiş olmalıdır. Bu hesaplanan rakamın altında denek içeren çalışmalardan elde edilecek pozitif veya negatif sonuçlar gerçeği yansıtmak yerine sadece rastlantısal değişkenliğe (random variation) bağlı olabilir. Rastlantısal değişkenliği açıklamak için iyi bir örnek 6 yüzlü zardır. Her yüzünde birden altıya kadar rakamlardan birisi bulan ve atıldığında her rakamın üste gelme olasılığı eşit olan hilesiz bir zar bir kez atıldığında her hangi bir rakamın gelme olasılığı  $1/6$ 'dır (yaklaşık %16.67). Ancak zarı sadece 6 kez attığımızda her zaman bir kez 1, bir kez 2, bir kez 3....bir kez 6 gelmesi olasılığı çok düşüktür. Zarı ancak bir çok kez arka arkaya attığımızda (yaklaşık 1000 kez mesela) toplamda 166 kez 1, 167 kez 2, 166 kez 3.... 166 kez 6 gelmesi daha muhtemeldir ki bu durumda bile rakamların geliş sıklıkları arasında küçük farklılıklar olacaktır. Atış sayımız arttıkça rakamların geliş sıklıkları doğal olasılığı olan %16.67'ye yaklaşacaktır. Az sayıda denek içeren çalışmalar yaparak gözlenen sonuçlardan genellemeye gitmek bir zarı 6 kez atıp her bir rakamın gelme sıklığı hakkında genelleme yapmaya benzer. Bu nedenle yeterli güce sahip olmayan araştırmaların sonuçlarına dikkatle bakmak ve rastlantısal değişkenliğin yansımaları olabileceğini akıldan çıkartmamak gerekir.

Örneklem hacmi hesabının yanında sonuçların karşılaştırılmasında kullanılan istatistik yöntemler açık bir şekilde belirtilmiş olmalıdır. Verilerin türlerine göre uygun hipotez testlerinin seçilip seçilmediği okuyucu tarafından kontrol edilmelidir.

## SONUÇLAR

*Araştırma sonuçları değerlendirilirken ilk bakılması gereken noktalar çalışmanın materyal ve metod bölümünde tarif edildiği şekilde tamamlanıp tamamlanmadığı ve katılan her bir hastaya ait sonuçların değerlendirmeye alınıp alınmadığıdır.*

Çalışmaya katılan ve randomize edilen her hastanın çalışma protokolünü tamamlayıp tamamlamamasına bakılmadan son analize dahil edilmesine "Intention to treat" prensibi denir (MOHER). Çalışmaya katılan hastalardan protokolü tamamlamayanların neden tamamlamadığının net bir şekilde açıklanması gereklidir. Deneklerin bir kısmı çalışma ilaçları veya uygulamalarının istenmeyen etkileri nedeniyle protokolü terk etmiş olabilir. Bu durumda benzer bir sorun söz konusu protokolü uygulamak istediğimizde bizim de başımıza gelebilir. Başka bir durum da araştırmacıların sonuçları bekledikleri veya istedikleri gibi olmayan bazı hastaları son analizden çeşitli bahanelerle hariç tutarak sonuçları etkileyebilmeleridir (post randomization selection). Bu durumu engellemek için çalışmaya katılıp randomize edilen her hasta son analize randomize edildiği grupta dahil edilmeli yani sonuçlar "Intention to treat" prensibine göre hesaplanmalıdır. Sonuçlar değerlendirilirken mutlaka paydalara dikkat edilmeli ve bu prensibe uyulduğundan emin olunmalıdır.

Sonuç ölçütlerinin birimleri net olarak belirtilmiş olmalıdır.

Kaliteli bir makalede araştırmaya katılan hastaların klinik uygulama için önemli olacak tüm sonuçları verilmiş olmalıdır. Örneğin, infertilite araştırmalarında sadece gebelik oranları değil, abortus oranları, çoğul gebelik oranları, yardımla üreme teknikleri ile ilgili bir çalışmaysa embryo implantasyon oranları gibi karar verme sürecimizi etkileyebilecek tüm sonuç ölçütleri açıkça belirtilmiş olmalıdır.

Sonuçlar bölümünde protokolün uygulanması sırasında deneklerde görülen istenmeyen/yan etkiler, komplikasyonlar açıklıkla belirtilmiş, okuyucunun kendi kendine

hesap yapmasına ihtiyaç bırakılmamış olmalıdır. Hiç istenmeyen etki gözlenmediyse bu durum açıkça bir cümleyle belirtilmiş olmalıdır.

## TARTIŞMA

*Tartışma bölümü araştırmanın temel bulgularının kısaca net olarak özetlendiği ve gözlenen sonuçların literatürdeki diğer veriler de göz önüne alınarak yorumlandığı bölümdür. Bu bölümde elde edilen sonuçların olası biyolojik açıklamaları ve hem klinik uygulamayı hem de gelecek araştırmaları nasıl yönlendireceği hakkında yazarlar düşüncelerini belirtir. Daha basit bir deyişle bulguların ne anlama geldiğinin açıklandığı bölümdür.*

Tartışma bir makalenin en subjektif bölümüdür. Okuyucu olarak dikkat edilmesi gereken noktalardan birisi yazarların çıkarımlarının sonuçlar bölümündeki bulgularla tutarlı olup olmadığıdır. Yazarların yorumları değerlendirilirken materyal ve metotta belirtilmiş olması gereken örneklem hacmi boyutu, çalışmanın istatistiksel gücü ve sonuçların mutlak değerleri göz önünde alınmalıdır. İncelenen gruplar arasında istatistiksel olarak anlamlı bir fark gözlenmesinin her zaman klinik olarak anlamlı olmayabileceği unutulmamalıdır. Diğer yandan az sayıda denek içeren bir çalışmada sonuçlar arasında mutlak değer olarak dikkat çekici bir fark olduğu halde istatistiksel anlamlılığa ulaşmamış olabilir. Bu durum yüksek beta hata düzeyine bağlı olabileceği gibi az sayıda denek nedeniyle rastlantısal değişkenliğe de bağlı olabilir ve gerçek etkinin ne olduğunu anlamak maalesef mümkün değildir.

Bir çok makalede yazarlar karşı konulamaz biçimde sonuçlarını bulunan rakamların yansıtıldığından öte bir

iyimserlikle yorumlama eğilimindedir. Yazarların sözel yorumlarıyla yetinmek yerine rakamların ne anlattığını değerlendirdikten sonra kendi yorumumuzu yapmamız gerekir.

Sonuçların literatürle kıyaslandığı durumlarda yazarların kendi bulgularını destekleyen yayınların yanında farklı sonuçlar bildiren yazıları da değerlendirmeye alıp almadığına dikkat edilmelidir. Gerekli görüldüğünde konu hakkında yayımlanan diğer makalelerin en azından özetlerine basit bir pubmed taramasıyla kolaylıkla ulaşılabileceği akılda bulundurulmalı ve bunu yapmaya üşenilmemelidir.

## KAYNAKLAR

1. Altman, D.G., The scandal of poor medical research. *BMJ*, 1994; 308(6924): p. 283- 4.
2. Arce, J.C., A. Nyboe Andersen, and J. Collins, Resolving methodological and clinical issues in the design of efficacy trials in assisted reproductive technologies: a mini-review. *Hum Reprod*, 2005; 20(7): p. 1757- 71.
3. Min, J. K., S. A. Breheny, V. MacLachlan, and D. L. Healy, What is the most relevant standard of success in assisted reproduction? The singleton, term gestation, live birth rate per cycle initiated: the BESST endpoint for assisted reproduction. *Hum Reprod*, 2004; 19(1): p. 3- 7.
4. Moher, D., K.F. Schulz, and D.G. Altman, The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*, 2001; 357(9263): p. 1191- 4.
5. Daya, S., The Placebo Effect. *Evidence Based Obstetrics and Gynecology*, 2000; 2(1): p. 1.
6. McCarney, R., J. Warner, S. Iliffe, R. van Haselen, M. Griffin, and P. Fisher, The Hawthorne Effect: a randomised, controlled trial. *BMC Med Res Methodol*, 2007; 7: p. 30.