

# SİSTEMATİK DERLEMELERİN KRİTİK ANALİZİ

Barış ATA, Bülent URMAN

Amerikan Hastanesi Kadın Sağlığı Merkezi, Tüp Bebek Ünitesi, İstanbul

## ÖZET

*Sistemik derlemeler bir araştırma sorusuna cevap arayan tüm randomize kontrolü çalışmaları bir araya getirerek bulguların sentezlenmesini sağlayan sekonder araştırma projeleridir. Bir derlemenin "sistemik" olarak tanımlanabilmesi için derlemeye katılacak çalışmaların belirlenmesi, seçilmesi, verilerin çıkartılması ve sentezlenmesi süreçlerinin sistemik, şeffaf ve tekrarlanabilir olması gereklidir. Doğru şekilde hazırlanan sistemik derlemeler klinik uygulamaların kanıta dayalı tıp ışığında yönlendirilmesinde en güvenilir kanıtları sağlamaktadır. Diğer yandan, sistemik derlemeler de insan emeğine dayanmaktadır ve küçük ölçekli birkaç çalışmayı bir araya getiren bir sistemik derlemenin sonuçlarıyla daha sonra gerçekleştirilen büyük ölçekli bir randomize çalışmanın ya da aynı soruya cevap arayan "sistemik derleme"lerin sonuçları da farklılık gösterebilmektedir. Bu makalede sistemik derlemelerin hazırlanma aşamaları ve sonuçlarının geçerliliğini ve genellenebilirliğini tehdit eden faktörler incelenerek, okuyucuların kendi yorumlarını yapmalarına yardımcı olmak amaçlanmıştır.*

**Anahtar kelimeler:** eleştirel analiz, kritik analiz, randomize kontrollü çalışma, sistemik derleme

*Türk Jinekoloji ve Obstetrik Derneği Dergisi, (TJOD Derg), 2008; Cilt: 5 Sayı: 4 Sayfa: 233- 40*

## SUMMARY

### CRITICAL APPRAISAL OF SYSTEMATIC REVIEWS

*Systematic reviews are secondary research projects that compile all the randomized trials addressing a particular question of interest. The processes and the procedures used to identify and select studies, extract and synthesize data must be standardized, transparent, and reproducible for a review to be classed as systematic. Well conducted, comprehensive systematic reviews of clinical trials provide the most valid evidence for directing the clinical practice in the light of evidence based medicine. However, systematic reviews rely on human endeavour, and it is known that results of a systematic review of relatively smaller studies may be contradictory to the results of a later single large randomized controlled trial, or two systematic reviews addressing the same question may provide different answers. The key stages of systematic reviews and factors threatening the validity of results of systematic reviews have been discussed in this article.*

**Key words:** critical appraisal, randomized controlled trial, systematic reviews

*Journal of Turkish Society of Obstetric and Gynecology, (J Turk Soc Obstet Gynecol), 2008; Vol: 5 Issue: 4 Pages: 233- 40*

---

**Yazışma adresi:** Uzm. Dr. Barış Ata: Amerikan Hastanesi Tüp Bebek Merkezi, Güzelbahçe sokak no 20 Nişantaşı, 34365 İstanbul, Türkiye  
Tel.: (0532) 744 74 16  
e-posta: barisata@hotmail.com

Alındığı tarih: 27.07.2008, revizyon sonrası alınma: 27.07.2008, kabul tarihi: 18.09.2008

## SİSTEMATİK DERLEMELER

### Sistemik derleme nedir?

Ne kadar iyi tasarlanmış ve yürütülmüş olursa olsun tek bir randomize klinik araştırmanın sonuçlarından genelleme yaparken yine de dikkatli olmak gereklidir. Bir araştırmanın sonuçlarının genellenebilirliği örneklem hacmi büyüklüğü ve katılımcıların çeşitliliği ile paralel olarak artmakla beraber çoğu zaman bir tedavi yöntemi ile ilgili birden fazla araştırma vardır ve bu araştırmaların sağladığı kanıtların biraraya getirilmesi daha çok genellenebilen ve klinik uygulamamızı yönlendirmekte daha güvenle kullanabileceğimiz daha güçlü kanıtlar sağlamaktadır. Bir tedavinin etkinliği hakkında en güvenilir kanıtı sağlayan araç olan kapsamlı sistemik derlemeler klinisyenler ve araştırmacılar için en değerli kaynaklardır.

Derleme makaleler genel olarak ikiye ayrılmaktadır; klasik akademik derlemeler ve sistemik derlemeler. Klasik akademik derlemeler araştırılan konu hakkında yazarların çeşitli kitap, dergi ve bildirilerden kendi görüşlerini destekleyen ve desteklemeyen sonuç ve yorumları biraraya getirerek nerdeyse bir düz yazı halinde sundukları derleme makaleleridir. Diğer yandan sistemik derlemeler sistemik ve rastlantısal hataları (bias ve random error) minimize etmek amacıyla belgelenmiş sistemik bir şekilde yapılan derlemelerdir<sup>(1)</sup>. Bir derlemenin “sistemik” olarak sınıflanabilmesi için en önemli kriter derlemeye katılacak araştırmaların belirlenmesi ve seçimi, seçilen araştırmalardan verilerin çıkartılması ve biraraya getirilerek sentezlenmesi sürecindeki işlemlerin standart ve şeffaf olmasıdır.

Sistemik derlemeler iki amaca hizmet ederler: 1. Belirli bir soruya cevap arayan mevcut TÜM randomize çalışmaların belirlenmesi ve özetlenmesi, 2. Bu çalışmaların sonuçlarının sentezlenmesi<sup>(2)</sup>. Doğru yapılmış bir sistemik derlemeyle biraraya getirilen ve sentezlenen bu sonuçlar araştırma sorusuna o an için geçerli en güvenilir cevabı verecektir.

### Sistemik derleme nasıl yapılır?

Sistemik derleme bir araştırma projesidir ve aynı klinik araştırmalar gibi önceden hazırlanmış bir protokol dahilinde yapılmalıdır. Bir sistemik derlemenin hazırlanma aşamaları aşağıdaki şekilde özetlenebilir<sup>(3)</sup>:

1. Araştırma sorusunun oluşturulması, katılım ve çıkarım kriterlerinin (inclusion and exclusion

criteria) tanımlanması.

2. Bu soruya cevap arayan çalışmaların literatür taraması ile tespit edilmesi.
3. Tanımlanmış olan katılım ve çıkarım kriterleri uygulanarak tespit edilen araştırmalardan derlemeye katılacakların belirlenmesi.
4. Derlemeye katılan primer araştırmaların özetlenmesi ve gerekli verilerin toplanması.
5. Çalışmaların kalitesinin değerlendirilmesi.
6. Verilerin analizi ve sentezi.
7. Sonuçların yorumlanması.

Bu noktada Sistemik Derleme ve Meta-Analiz arasındaki farkı vurgulamak yerinde olacaktır. Sistemik derleme terimi çalışmaların tespiti ve belirlenmesi, verilerin toplanması, sonuçların sentezi dahil olmak üzere derleme sürecinin tamamını tanımlarken, meta-analiz derlenen çalışmaların sonuçlarını biraraya getirerek sentezlemek için kullanılan bir istatistiksel tekniktir. Bazen çalışmaların sonuçlarının heterojen olması geçerli bir meta-analiz yapılmasına engel olabilir, ve meta-analiz içermeyen bir sistemik derleme yapılabilir. Bazen de sadece veriler toplanarak bir meta-analiz yapılabilir ve sistemik derlemenin diğer bölümleri yazılmayabilir. Bu iki terim sıklıkla birbirinin yerine kullanılmakta olsa da farklı kavramları tanımlamaktadırlar ve birbirlerinin yerine geçmezler.

### Bir sistemik derlemenin eleştirel analizi

Küçük ölçekli birkaç çalışmayı bir araya getiren bir sistemik derlemenin sonuçlarıyla daha sonra gerçekleştirilen büyük ölçekli bir randomize çalışmanın sonuçları çelişebilmektedir, benzer şekilde aynı soruya cevap arayan “sistemik derleme”lerin sonuçları da farklılık gösterebilmektedir<sup>(4)</sup>. Bu noktada “sistemik” derleme hazırlanışı şeffaf bir şekilde ortaya konulmuş kendi protokolü dahilinde yapılan derlemeyi tanımlamaktadır. Aynı soruyu cevaplamak amacıyla hazırlanan farklı sistemik derlemelerin yazarları farklı kaynakları taramış, farklı katılım ve çıkarım kriterleri kullanmış, farklı çalışmaları derlemelerine dahil etmiş, farklı istatistiksel yöntemleri kullanmış olabilirler. Bu nedenle okuyucunun “sistemik” olarak tanımlanmış bir derlemeyi değerlendirirken derlemeyi yapanların sonuçların geçerliliğini (internal validity) ve genellenebilirliğini (external validity) maksimize etmek için izledikleri yöntemin kalitesini değerlendirerek kendi yorumunu yapabilmesi gereklidir.

Sistemantik derlemelerin eleştirel analizi birçok noktada klinik araştırma makalelerinin kritik analizi ile benzerlik göstermekle beraber derlemeler için ilave birtakım potansiyel hata kaynakları mevcuttur ve bunların da ayrıca gözden geçirilmesi şarttır. Bu hata kaynakları yayınlanma farklılığı (publication bias) ve bildirilme farklılığı (reporting bias) olup aşağıda detaylı olarak ele alınacaktır.

SDlerin eleştirel analizi temel olarak orjinal araştırma makalelerinin eleştirel analizi ile benzerlik gösterir. Cevaplanması gereken sorular şunlardır: 1. Araştırılan soru nedir? 2. Sonuçlar güvenilir midir? (internal validity) 3. Sonuçlar/yorumlar uygulamaya nasıl yansıtılabilir/genellenebilir? (external validity) 4. Katılan çalışmaların ve derlemenin sonuçları ne anlama gelmektedir? Klinik araştırma makaleleri için bu soruların nasıl irdeleneceği Derginin önceki sayısında yayınlanan makalemizin konusu olduğundan burada sadece kısaca ele alınacak, ancak SDlerin kritik analizinde farklılık gösteren noktalar vurgulanacaktır (5).

#### ***Derlemenin cevaplama çalıştığı soru nedir?***

Yukarıda belirtildiği gibi SDler birer araştırma projesidir ve her klinik araştırma projesinde olduğu gibi araştırma sorusu spesifik olarak belirlenmiş olmalıdır. Soruda popülasyon, girişim/televi, kontrol girişim/televisi, sonuç ölçütleri ve gerekli olduğunda zaman açık şekilde belirlenmiş ve ifade edilmiş olmalıdır. Sorusu net olmayan bir araştırmanın sonucunda bulunacak cevaplarının da net olamayacağı açıktır.

#### ***Sonuçlara güvenilebilir mi? (Internal validity)***

SDler retrospektif gözlemsel araştırmalardır ve derlemeye dahil edilen primer araştırmalar gibi sonuçların güvenilirliği çeşitli faktörler tarafından tehdit edilmektedir. SDler için söz konusu olan sistemantik hata (bias) kaynakları üç başlıkta toplanabilir:

1. Yayınlanma farklılığı (Publication bias).
2. Seçim yanlılığı (Selection bias).
3. Derlemeye dahil edilen primer araştırmaların kalitesindeki değişkenlik.

#### ***Yayınlanma Farklılığı ve Tarama Stratejisi***

Pozitif bulgu bildiren araştırma makalelerinin yayınlanma olasılığı daha yüksektir<sup>(6)</sup>. Ek olarak bu araştırmaların İngilizce bir dergide yayınlanma olasılığı daha yüksek, tamamlanmasından yayınlanmasına kadar

geçen süre daha kısa (time lag bias) ve bu makalelere atıfta bulunulması olasılığı daha yüksektir<sup>(7)</sup>. Bu gerçeğin arkasında yatan iki neden vardır; dergi editörlerinin negatif sonuç bildiren çalışmaları yayınlamaktan pozitif sonuç bildirenleri yayınlamaya daha istekli olması ve araştırmacıların pozitif bulgu veren çalışmalarını yayınlamakta daha istekli ve kararlı olup daha hızlı davranmasıdır. Pozitif bulguların negatif bulgulardan çok yayınlıyor olmasının bir nedeni de araştırmacıların çalışmalarında elde ettikleri sonuçlar içerisinden pozitif olanları bildirirken negatif sonuçları göz ardı etmeye olan eğilimleridir (selective reporting of outcomes).

Bir tedavinin etkinliğini değerlendiren çalışmalardan pozitif bulgu verenlerin daha çok ve daha hızlı yayınlanması, haliyle bu makalelerin literatürde birikerek negatif sonuç veren çalışmalar yayınlanana kadar daha çok atıf alıp daha kolay erişilebilir olması, konuyu inceleyen bir kişinin literatürü taradığında tedavinin etkinliğinin gerçekte olduğundan daha yüksek olduğunu düşünmesine neden olmaktadır. Aynı şekilde bu çalışmaları içeren bir derlemede çalışmaların verilerini bir araya getiren bir metaanaliz de bu etkiyi daha da abartılı şekilde ortaya koyabilmektedir. Majör tıbbi dergilerin editörleri bu çok önemli sorunun önüne geçmek amacıyla araştırma sonuçlarının yayınlanabilmesi için yürütülecek araştırma projelerinin çalışma başlamadan önce herkesin erişimine açık online veritabanlarına kayıt ettirilmesi zorunluluğunu getirmiştir. Bu kural giderek artan sayıda dergi tarafından uygulanmaktadır. Önceden kayıt yaptırılırken diğer bilgilerin yanında çalışmanın amacı, çalışmaya katılacak popülasyon, incelenecek hastalık/durum, araştırılacak tedavi/girişim ve belki de en önemlisi değerlendirilecek sonuç ölçütleri bildirilmektedir. Böylece araştırma tamamlandığında önceden belirlenmemiş ancak araştırmacılar tarafından ilgi çekici bulunan, belki de tesadüfi olan pozitif bulguların bildirilmesine engel olurken diğer yandan çalışma protokolünde baştan belirlenmiş olduğu halde pozitif bulgu vermeyen sonuç ölçütlerinin bildirilmemesine engel olunması amaçlanmaktadır.

Kayıt sisteminin bir faydası da SD hazırlayacak araştırmacıların bir konu üzerinde sadece yayınlanmış olan çalışmaları değil, tamamlanmış ama yayınlan(a) mamış çalışmaları ve sürmekte olan çalışmaları da belirleyebilmek gerekrise araştırmacılar ile temasa geçerek SDye dahil edebilmelerine olanak sağlamasıdır.

Araştırma kaydı yapılan veritabanlarını daha detaylı incelemek isteyen okuyucular yaygın olarak kullanılan veritabanlarının websitelerine şu adreslerden ulaşabilirler: [www.clinicaltrials.gov](http://www.clinicaltrials.gov), [www.isrctn.org](http://www.isrctn.org) ve [www.who.int/ictrp/data\\_set/en](http://www.who.int/ictrp/data_set/en).

Yayınlanma farklılığı hakkında belirtilenler göz önüne alındığında bir SDye değerlendirirken, araştırmacıların SDye katılacak araştırmaları nasıl tespit ettiklerinin dikkatle incelenmesi gerektiği açıktır. Online kayıt sistemlerine ek olarak taranan literatür veri tabanlarına da dikkat edilmesi gereklidir. En yaygın kullanılan veritabanı olan MEDLINE araştırılan konuya göre değişmekle beraber yayınlanmış tüm randomize çalışmaların %30 - 80'ini içermekte olup, çoğunlukla İngilizce yayınlanan dergileri indekslemektedir<sup>(8)</sup>. Sadece MEDLINE ile kısıtlı bir taramanın mevcut araştırmaların tamamını kapsamayacağı ve pozitif sonuç bildirenlerin bulunma olasılığını artıracığı açıktır. Yaygın kullanılan bir diğer veritabanı olan EMBASE ile MEDLINE'in indeksledikleri dergilerin örtüşme oranı %40 civarındadır. Yani MEDLINE'a ek olarak sadece EMBASE'in de taranması çok sayıda ilave randomize araştırmanın bulunmasını sağlayacaktır. Günümüze randomize kontrollü çalışmaları belirlemek için ilk sırada ve mutlaka taranması gereken en kapsamlı kaynak The Cochrane Controlled Trials Register'dır (CCTR). Cochrane Collaboration mevcut tüm veritabanlarını anahtar kelimeler, çalışma tasarımı özellikleri ve araştırma özetlerini içerecek şekilde taramakta ve veritabanlarının yöneticileri ile işbirliği içerisinde belirlediği randomize kontrollü çalışmaların "randomize kontrollü çalışma" şeklinde etiketlenmesini sağlayarak daha sonra veritabanlarını tarayacak araştırmacılar tarafından bulunmasını kolaylaştırmanın yanında bu çalışmaları ayrıca konu başlıklarına göre CCTR'da kaydetmektedir.

Altın standart olarak kabul edilen Cochrane Sistematik Derlemeleri kalitesinde bir derleme için yaygın kullanılan veri tabanları dışında araştırılan konu ile ilgili indekslenmeyen dergilerden gerekli görülenlerin, kongre ve toplantı bildirimlerini içeren özet kitapçıklarının, hatta konu ile ilgili eğitim veren bazı önemli kurumların kütüphanelerindeki veya ulusal indekslerdeki tez çalışmalarının da manuel taranması gereklidir. Yapılan bir araştırmada kongre özet kitaplarında yayınlanan bildiriler ve the Johns Hopkins School of Public Health gibi major eğitim kurumlarındaki tez çalışmaları içerisindeki randomize kontrollü çalışmaların yaklaşık

%50 - 70'inin indekslenen dergilerde tam metin yayın aşamasına ulaştığı görülmüştür<sup>(9-15)</sup>. Ek olarak bu çalışmalardan sonunda yayın aşamasına ulaşanların diğer makalelere kıyasla anlamlı olarak daha çok negatif sonuç içerdiği gözlenmiştir<sup>(16)</sup>.

Bu kadar detaylı bir taramanın çok zaman, emek ve kaynak gerektirdiği açıktır. Gerçek dünyada tabii ki araştırmacılar araştırma stratejilerini ellerindeki kaynaklara göre belirlemek zorundadır, ancak bu durum okuyucunun makaleyi okurken ki değerlendirmesini etkilemeyecektir. Yukarıda sayılan kayıt veritabanları, konferans özetleri, dergilerde yayınlanmamış tez çalışmaları genel olarak "gri literatür" olarak isimlendirilir ve gri literatürdeki araştırmaların indeksli dergilerde yayınlanan araştırmalara kıyasla genel olarak daha düşük kalitede olduğu düşünülmektedir. Yine de böyle bir önyargıyla gri literatürü tamamen gözardı etmektense tarayıp gri literatürden bulunan çalışmaları da SD protokolünde önceden belirlenmiş olan kriterlere göre elemek en doğru yöntemdir.

Kapsamlı bir SD yapılabilmesi için birden fazla kaynak taranmalı, tarma süreci sistematik, planlı, önceden belirlenen bir protokole göre yapılmış, araştırma sorusuna uygun, mevcut kaynaklar ve zaman sınırlaması içerisinde olabildiğince kapsamlı ve şeffaf (okuyucuların ne yapıldığını görebileceği ve tarama süreci başka araştırmacılar tarafından tekrarlanabilir ve genişletilebilir şekilde) olmalıdır.

Taranan kaynaklara ek olarak taramada kullanılan anahtar kelimelerin seçimi ve dökümantasyonu da çok önemlidir. Kapsamlı bir taramada araştırma sorusunun her bir bileşeni (popülasyon, hastalık/durum, araştırılan tedavi/girişim, kontrol tedavi/girişimi, sonuç ölçütleri ve gerektiğinde ilgilenilen zaman dilimi) ile ilgili anahtar kelimeler kullanılmış olmalıdır. Anahtar kelimeler seçilirken sadece medikal terimler değil paramedikal karşılıkları da (infertilite ve kısırlık gibi) kullanılmış olmalıdır. Araştırma konusu ilaçları içeriyorsa hem jenerik hem de marka isimleri kullanılmalıdır. Anahtar kelimelerin olası tüm farklı yazım şekilleri taranmış olmalıdır. SDler konusunda "kalite standardı" olarak değerlendirilen Cochrane Sistematik Derlemeleri incelendiğinde taranan anahtar kelimelerin bir liste halinde hatta bazılarında araştırma sorusunun bileşenlerinden oluşan başlıklar altında bir tabloda verildiği görülebilir<sup>(17,18)</sup>.

Okuyucunun tarama işleminin kalitesini değerlendirebilmesi için Tablo I'de belirtilen maddelerin herbirisinin

yöntem bölümünde açıkça belirtilmiş olması ve okuyucu tarafında değerlendirilmesi gereklidir. Tarama yöntemi ne kadar kapsamlı olsa da bulunan ve derlemeye katılan çalışmaların gerçekten tüm olası sonuçları içerdiğinden emin olmak mümkün olmayacaktır. Gereçeğe ne kadar yaklaşılabildiğini ve yayınlanma farklılığının gerçekten mevcut olup olmadığını değerlendirmek için bazı istatistik yöntemler mevcuttur. Bu yöntemlerden en yaygın kullanılan ve en basitlerinden birisi “funnel plot”tır. Terimin Türkçe karşılığı the International Statistical Institute sözlüğünde “huni grafiği” olarak verilmektedir. Huni grafiği basit bir tanımla her bir çalışmada gözlenen tedavi etkinliğinin çalışmanın örneklem hacmi büyüklüğüne göre dağılımını gösteren bir saçılım grafiğidir (scatterplot). Örneklem hacmi büyüklüğünün kendisi kullanılabilir de matematiksel nedenlerden ötürü genellikle tedavi etkinliği tahmininin (estimate of treatment effect) standart hatası veya varyansı tercih edilmektedir. Yayınlanma farklılığı olmadığında grafiğin Şekil 1’deki gibi simetrik ters bir huni şeklinde olması beklenir. Huni grafiğinin altında yatan prensip daha büyük örneklem hacmine sahip çalışmalarda gerçek tedavi etkinliği tahmininin daha kesin olacağıdır. Her bir noktanın bir çalışmayı temsil ettiği Şekil 1’deki grafiğin alt kısmında daha küçük örneklem hacimli çalışmalarda gözlenen tahminler daha geniş bir dağılım gösterirken, üst kısımlardaki daha büyük örneklem hacimli çalışmalarda gözlenen tahminler giderek daha dar dağılım göstermektedir<sup>(19)</sup>. Yayınlanma farklılığı varlığında grafiğin alt köşelerinden bir tanesindeki noktalar bulunmayacaktır. Bu durumda Şekil 1’de, boş daireler grafiğin alt kısımlarında geniş yayılım gösteren küçük örneklem hacimli bu nedenle abartılı odds oranları veren ve negatif sonuç bildiren çalışmalar olup hiç yayınlanamamış olacaktı ve Şekil 1’de sadece dolu siyah daireler olacaktır. Halbuki gerçekte bir tedavinin bir hastalıkta benzer popülasyonlar üzerinde bir tek etkinliği olacaktır, ve yapılan çeşitli çalışmalarda gözlenen etkinlik tahminleri de doğal varyasyon dahilinde bu gerçek değer etrafında normal simetrik dağılım gösterecektir. Her çalışmada gözlenen etkinlik tahminleri de çalışmaların örneklem hacimlerine göre Y eksenine sıralanarak grafiğe ikinci boyutu katıldığında huni eğrisi elde edilmektedir.

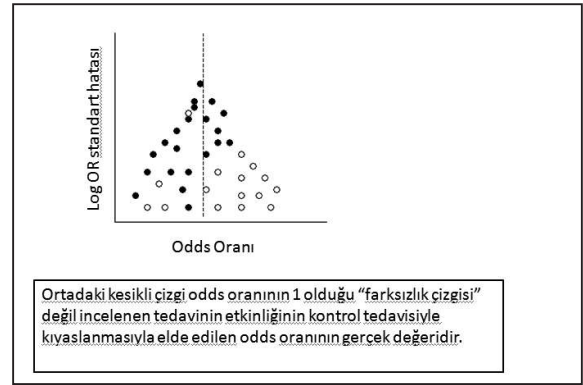
**Tablo I: Sistemantik derlemeler için tarama sürecinin tanımlanması**

#### Taranan Elektronik Veritabanları

- Taranan veritabanının ismi (MEDLINE, EMBASE, LILACS vs)
- Tarama yapılan servis sağlayıcının (host) ismi
- Taramanın yapıldığı tarih
- Tarama sürecinin kapsadığı yıllar
- Uygulanan tarama yöntemi (kullanılan tüm anahtar kelimeler dahil olmak üzere)
- Varsa dil kısıtlaması (sadece İngilizce yayınlar gibi)

#### Manuel Taranan Dergiler

- Dergilerin tam isimleri
- Taranan yıl ve sayılar (eksikler varsa belirtilmek üzere)
- Yayınlanmamış araştırmaları belirlemek için kullanılan yöntemler
- Taranan kongre özet kitapları
- Danışılan uzmanlar
- Referans listesi taranan makalelerin listesi
- Taranan gri literatür veritabanları



**Şekil 1: Yayınlanma farklılığı olmadığında gözlenmesi beklenen hipotetik huni grafiği**

Huni eğrilerine bakılarak tespit edilen ve derlemeye katılan çalışmalar için yayınlanma farklılığı olup olmadığı kolaylıkla değerlendirilebilmektedir. Asimetrik huni eğrisi oluşmasının tek nedeni yayınlanma farklılığı olmayabilir, olası her çalışma henüz yapılmamış olduğundan normal dağılım gözlenemiyor olabilir (derlemenin örneklem hacmi küçük olabilir), küçük örneklem hacimli çalışmaların metodolojik kalitesindeki sorunlardan dolayı yansıttıkları etkinlik gerçek dışı olabilir, artefakt bulunabilir. Asimetrik huni eğrisinin nedeni ne olursa olsun bu durumda bilinmesi gereken şey mevcut çalışmaların sonuçlarını bir araya getiren bir metaanaliz yapılması halinde hesaplanacak tedavi etkinliğinin abartılı olacağıdır<sup>(20)</sup>.



### ***Derlemeye katılacak çalışmaların seçimi ve Seçim yanlılığı***

Tarama sonucunda tespit edilen çalışmalar değerlendirilerek elenecektir. Bu eleme işlemi SD protokolünde önceden belirlenmiş olan kriterlere göre objektif şekilde gerçekleştirilmiş olmalıdır. Katılım kriterleri araştırma sorusuna göre belirlenmiş olmalı ve sorunun her bir bileşenini değerlendirerek aşağıdaki noktaları net olarak belirlemelidir. Araştırma sorusunu oluşturan bileşenler Derginin önceki sayısındaki makalemizde ve bu makalenin önceki bölümlerinde belirtilmiş olup popülasyon, tedavi/girişim, kontrol tedavi/girişimi, sonuç ölçütleri ve zamandır<sup>(5)</sup>.

SDye katılacak çalışmaların nasıl bir popülasyonda yapılmış olması gerektiği, hangi girişim/tedaviyi incelemiş olması gerektiği, bu tedavi/girişimi nasıl bir kontrol grubuyla karşılaştırmış olması gerektiği, hangi sonuç ölçütlerinin değerlendirmiş olması gerektiği açıkça belirtilmiş olmalıdır. Çalışma seçiminde bir önemli faktör de çalışma tasarımıdır. SDler her zaman sadece randomize kontrollü çalışmaları içermezler. Ancak bir tedavinin etkinliğinin en iyi değerlendirildiği tasarımın randomize kontrollü çalışma olması nedeniyle en iyi SDler de sadece randomize kontrollü çalışmaları içeren SDlerdir. Sonuçta SDlerin kalitesi dahil edilen çalışmaların kalitesine bağlıdır.

Katılım ve çıkarım kriterlerini net bir şekilde belirlemek ve belirli bir çalışmanın da bu kriterlere uyup uymadığına karar vermek her zaman kolay değildir. Ne kadar planlı ve önceden hazırlanmış bir protokole göre yapılsa da sonuçta bütün bu işlemler insan faktörüne dayanmaktadır. Herhangi bir çalışmanın hazırlanan SDye katılmak için uygun olup olmadığı konusunda görüş ayrılıkları olabilmektedir. Objektif bir seçim sürecinde çalışmalar birden fazla yazar tarafından bağımsız olarak değerlendirilmiş ve daha sonra sonuçlar karşılaştırılarak mevcut fikir ayrılıkları uzlaşa arayarak giderilmiş olmalıdır. Çalışmaların elenmesi sürecinde en az iki araştırmacının rol almasının konu ile ilgili çalışmaların yanlılıkla elenmesi olasılığını azalttığı gözlenmiştir<sup>(21)</sup>. Seçim sürecinde yazarların değerlendirmelerini yaparken çalışmanın yayınlandığı derginin veya yazarlarının isminden etkilenmemesi için çalışmaları bu bilgilerin gizlendiği kopyalardan okuyarak değerlendirmiş olmaları ideal yöntemdir. Ancak bu yöntemin gerçek hayatta uygulanmasının zorluğunun ötesinde SDyi hazırlayan araştırmacılar araştırılan konuya duydukları ilgi ve deneyimleri nedeniyle çalışmaları sadece içeriğinden

de tanıyabilirler.

Tespit edilen çalışmalardan SDye dahil edilmeyenlerin bir listesi hariç tutulma nedenleriyle beraber okuyucuya sunulmalı ve çalışma seçiminin önceden belirlenen objektif kriterlere göre yapıp yapılmadığının okuyucu tarafından değerlendirilmesine olanak sağlanmış olmalıdır. Sadece SDye katılan çalışmaların tanıtılması kesinlikle yeterli değildir. Bütün bu detayların amacı SDyi hazırlayanların bilinçli veya bilinçsiz olarak mevcut çalışmaların içerisinden kendi görüşlerini destekleyen çalışmaları dahil ederken karşıt bulgu vermiş çalışmaları sadece sonuçlarına bakarak hariç tutmadıklarından emin olmaktır. Bazen tek bir çalışmanın katılması veya çıkartılması meta-analiz sonuçlarının yönünü değiştirebilmektedir.

### ***Derlemeye katılan primer araştırmaların kalitesindeki değişkenlik***

Yukarıda da belirtildiği gibi bir SDnin kalitesi en başta o SDye dahil edilen primer araştırmaların kalitesine bağlıdır. Diğer bir deyişle hammadde ne kadar kaliteliyse son ürün de ancak o kadar kaliteli olabilir. Klinik araştırmalar konusunda “kalite”yi tanımlamak zordur. Çalışma tasarımı, çalışmanın yürütülmesi, verilerin analizi, bulguların klinik önemi veya makalenin yazımı konularının herbirisinin kalitesi ayrı ayrı değerlendirilebilir. Bu bölümde SDlere katılacak çalışmaların kalitesini değerlendirirken yorum kolaylığı olması amacıyla kalite kavramını sonuçların geçerliliği (internal validity) ve genellenebilirliği (external validity) olmak üzere iki ana başlıkta ele alınması tercih edilmiştir. Genellenebilirlik (external validity) bir araştırmadan elde edilen sonuçların hangi koşullarda ve kimleri tedavi ederken yol gösterici olabileceği şeklinde tanımlanabilir. Örnek olarak hipogonadotropik hipogonadizm nedeniyle anovulatuvar olan bir kadına ovülasyon indüksiyonu yaparken protokol tercihi polikistik over sendromu (PKOS) nedeniyle anovulatuvar olan kadınlarda ovülasyon indüksiyonu sonuçlarını inceleyen çalışmalara dayandırılmaz, benzer bir popülasyonda yani hipogonadotropik hipogonadizimli hastalarda yapılan araştırma sonuçlarına bakarak karar verilmesi gerekir. PKOSlu kadınlarda ovülasyon indüksiyonunu değerlendiren çalışmaların genellenebilirliği PKOSlu kadınlarda sınırlıdır. Bu basit örneği zayıf PKOSlu hastalar için zayıf PKOSlu kadınlarda yapılan çalışmalara bakmak, Türkiye'deki hastaları tedavi ederken beyaz ırktan hastaları içeren yabancı çalışmalara bakmak gibi

genişletmek mümkündür.

Internal validity kavramı bir çalışma sonucunda gruplar arasında gözlenen farkların rastlantısal hata olasılığı dışında araştırılan tedavi/girişime bağlanabilmesi anlamına gelir. Internal validity bir araştırma sonuçlarının doğruluğunu veya başka bir deyişle geçerliliğini tanımlayan bir kavramdır. Bir araştırmanın sonuçlarının geçerliliğini tehdite eden çeşitli sistemantik hata kaynakları olabilir. Kısaca çalışma gruplarındaki hastaların özelliklerindeki farklılıklar (seçim yanlılığı, selection bias)(buradaki seçim yanlılığı sistemantik derleme için tespit edilen çalışmalardan derlemeye katılacak olan çalışmaların seçimindeki yanlılık değil, bir araştırmada katılımcıların çalışma gruplarına ayrımı sırasındaki seçim yanlılığıdır), gruplar arasında deneysel tedavi/girişim dışındaki bakımın farklı olması (treatment veya performance bias), sonuç ölçümlerinde yanlılık (detection veya ascertainment bias) ve protokolü tamamlamayan katılımcıların analize dahil edilme yöntemindeki yanlılık (attrition bias) bir klinik araştırmanın sonuçlarının geçerliliğini tehdit eden faktörler olarak sayılabilir. Bu faktörlerin herbirisi Derginin bir önceki sayısındaki makalemizde tartışılmış olduğundan burada sadece isimlerinin verilmesiyle yetinilecektir<sup>(5)</sup>.

Bir SDye dahil edilen araştırmaların kalitesini değerlendirirken yazarların dahil ettikleri ve hariç tuttukları çalışmaların genellenebilirliklerinin ve geçerliliklerinin okuyucu tarafından kontrol edilmesi gerekir. Bir araştırmanın sonuçlarının geçerliliği şüpheli olduğunda sonuçlardan kimler için genelleme yapılabileceği sorusu anlamını yitirmektedir. Ancak, bir SDyi okurken kolaylık olması için bu sıra tersine çevrilebilir. Öncelikle dahil edilen çalışmaların SDnin araştırma sorusuna cevap arayan çalışmalar olup olmadığını ve bu soruya cevap arayan tüm çalışmaların değerlendirmeye alınıp alınmadığına bakılması uygun olacaktır. Bu amaçla herbir araştırmanın katılımcı popülasyonunu, çalışma ve kontrol gruplarına uygulanan tedavi ve girişimleri ile değerlendirilen sonuç ölçütlerine bakılmalıdır.

Doğru çalışmaların SDye dahil edildiğini gördükten sonra bu çalışmaların herbirisinin sonuçlarının geçerliliğini değerlendirmek amacıyla herbir çalışmada katılımcıların çalışma ve kontrol gruplarına ayrımında randomizasyon kurallarına uyulup uyulmadığını, her iki gruba uygulanan tedavi/girişimler arasında tek farkın etkinliği araştırılan tedavi/girişim olup olmadığını,

çalışmanın yürütülmesi ve sonuçların değerlendirilmesi sırasında körlüme (masking veya blinding) uygulanıp uygulanmadığını, herbir çalışmadan veriler derlenirken Intention - to - treat prensibine uyulup uyulmadığının kontrol edilmesi gerekir.

Bu aşamalarda her zaman orjinal makalelerin tam metinlerinin bulunup değerlendirilmesi pratik nedenlerden ötürü beklenemez. Ancak kaliteli bir SDde yazarların katılan ve hariç tutulan çalışmalar için ayrı ayrı olmak üzere bütün bu konuları birer başlık altında birer ikişer cümleyle özetledikleri iki liste sunmuş olması idealdir. SDler için altın standart olarak kabul edilen Cochrane SDlerinde bu iki liste her zaman verilmektedir. Bu iki listeyi gözden geçirdikten sonra şüpheler varsa o zaman orjinal araştırma makalelerinin incelenmesi uygun olacaktır.

Bu makalede SDlerin hazırlanış sürecinin nasıl değerlendirileceği anlatılmıştır. Bir SD hazırlanırken araştırma sorusuna cevap arayan tüm randomize kontrollü çalışmaların tespit edilebilmesi için gerekli koşullar ve tespit edilen çalışmalardan SD için uygun olanların seçimi tartışılarak yazarların konu hakkında mevcut tüm verileri doğru şekilde biraraya getirip getirmediğinin değerlendirilmesi için dikkat edilmesi gerekli noktalar sıralanmıştır. Tespit edilen ve katılan çalışmaların geçerliliklerinin değerlendirilmesi konusu Derginin önceki sayısındaki Klinik Araştırma Makalelerinin Eleştirel Analizi başlıklı makalemizde daha geniş ele alındığından burada kısaca bahsedilmiştir<sup>(5)</sup>. Doğru çalışmaların biraraya getirilmesinden sonra herbir çalışmadan gerekli verilerin toplanması ve toplu halde değerlendirilmesinde kullanılan istatistik yöntemlerin özellikleri ayrı bir makalede ele alınacaktır.

## KAYNAKLAR

1. Chalmers, I. and Altman, D., Systematic Reviews. 1995, London: BMJ Books.
2. Egger, M., Rationale, potentials, and promise of systematic reviews. in: M. Egger, G.D. Smith, and D. Altman Eds Systematic reviews in healthcare: Meta-Analysis in context. London: BMJ Books, 2001:
3. Evans, S., Newman, M., and Thomas, S., Introduction to systematic reviews: Why do we need them and what do they do? in: D. Elbourne and K. Searle Eds Reporting and Reviewing Clinical Trials, PG Diploma/MSc in Clinical Trials. London: London School of Hygiene and Tropical Medicine, 2007:

4. LeLorier, J., Gregoire, G., Benhaddad, A., Lapierre, J., and Derderian, F., Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med*, 1997; 337(8): p. 536- 42.
5. Ata, B. and Urman, B., Klinik araştırma makalelerinin eleştirel analizi. *TJOD Dergisi*, 2008; 5(2): p. 83- 8.
6. Sterling, T., Rosenbaum, W., and Weinkam, J., Publication decisions revisited: the effect of the outcome of statistical tests on the decision to publish vice versa. *Am Stat*, 1995; 49: p. 108 - 12.
7. Egger, M., Dickersin, K., and Smith, G.D., Problems and limitations in conducting systematic reviews. in: M. Egger, G.D. Smith, and D. Altman Eds *Systematic reviews in healthcare Meta-analysis in context*. London: BMJ Books, 2007: 48- 9.
8. Dickersin, K., Scherer, R., and Lefebvre, C., Identifying relevant studies for systematic reviews. *BMJ*, 1994; 309: p. 1286 - 91.
9. Donaldson, I.J. and Cresswell, P.A., Dissemination of the work of public health medicine trainees in peer-reviewed publications: an unfulfilled potential. *Public Health*, 1996; 110(1): p. 61-3.
10. Scherer, R.W., Dickersin, K., and Langenberg, P., Full publication of results initially presented in abstracts. A meta-analysis. *JAMA*, 1994; 272(2): p. 158- 62.
11. Egger, M., Juni, P., Bartlett, C., Holenstein, F., and Sterne, J., How important are comprehensive searches and assessment of trial quality in systematic reviews? Empirical Study. *Health Technol Assess*, 2003; 7(1).
12. Goldman, L. and Loscalzo, A., Fate of cardiology research originally published in abstract form. *N Engl J Med*, 1980; 303(5): p. 255- 9.
13. Meranze, J., Ellison, N., and Greenhow, D.E., Publications resulting from anesthesia meeting abstracts. *Anesth Analg*, 1982; 61(5): p. 445- 8.
14. De Bellefeuille, C., Morrison, C.A., and Tannock, I.F., The fate of abstracts submitted to a cancer meeting: factors which influence presentation and subsequent publication. *Ann Oncol*, 1992; 3(3): p. 187- 91.
15. Juzych, M.S., Shin, D.H., Coffey, J.B., Parrow, K.A., Tsai, C.S., and Briggs, K.S., Pattern of publication of ophthalmic abstracts in peer-reviewed journals. *Ophthalmology*, 1991; 98(4): p. 553- 6.
16. McAuley, L., Pham, B., Tugwell, P., and Moher, D., Does the inclusion of grey literature influence estimates of intervention effectiveness reported in metaanalyses? *Lancet*, 2000; 356: p. 1228 - 1231.
17. Mochtar, M.H., Van der, V., Ziech, M., and van Wely, M., Recombinant Luteinizing Hormone (rLH) for controlled ovarian hyperstimulation in assisted reproductive cycles. *Cochrane Database Syst Rev*, 2007(2): p. CD005070.
18. Underhill, K., Operario, D., and Montgomery, P., Abstinence-only programs for HIV infection prevention in high-income countries. *Cochrane Database Syst Rev*, 2007(4): p. CD00 5421.
19. Deeks, J., Higgins, P., and Altman, D., Analysing and presenting results. in: J. Higgins and S. Green Eds *Cochrane Handbook of Systematic Reviews of Interventions*. 2005: 97- 166.
20. Egger, M., Davey Smith, G., Schneider, M., and Minder, C., Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 1997; 315: p. 629- 34.
21. Edwards, P., Clarke, M., DiGuseppi, C., Pratap, S., Roberts, I., and Wentz, R., Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. *Stat Med*, 2002; 21(11): p. 1635- 40.