



T.C.  
BURSA ULUDAĞ ÜNİVERSİTESİ  
SAĞLIK BİLİMLERİ ENSTİTÜSÜ  
TIP FAKÜLTESİ  
BİYOİSTATİSTİK ANABİLİM  
DALI



**META ANALİZİNDE YANLILIK DEĞERLENDİRME  
YÖNTEMLERİNİN KARŞILAŞTIRILMASI**

**Fisun KAŞKIR KESİN**

**(DOKTORA TEZİ)**

**BURSA-2022**

Fisun KAŞKIR KESİN

BİYOİSTATİSTİK ANABİLİM DALI DOKTORA TEZİ

2022



T.C.  
BURSA ULUDAĞ ÜNİVERSİTESİ  
SAĞLIK BİLİMLERİ ENSTİTÜSÜ  
TIP FAKÜLTESİ  
BİYOİSTATİSTİK ANABİLİM DALI



**META ANALİZİNDE YANLILIK DEĞERLENDİRME  
YÖNTEMLERİNİN KARŞILAŞTIRILMASI**

**Fisun KAŞKIR KESİN**

**(DOKTORA TEZİ)**

**DANIŞMAN:  
Prof.Dr. İlker ERCAN**

**BURSA-2022**

**T.C.  
BURSA ULUDAĞ ÜNİVERSİTESİ  
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**ETİK BEYANI**

Doktora tezi olarak sunduğum

“Meta Analizinde Yanlılık Değerlendirme Yöntemlerinin Karşılaştırılması” adlı çalışmanın, proje safhasından sonuçlanmasına kadar geçen bütün süreçlerde bilimsel etik kurallarına uygun bir şekilde hazırlandığını ve yararlandığım eserlerin kaynaklar bölümünde gösterilenlerden oluştuğunu belirtir ve beyan ederim.

**Fisun KAŞKIR KESİN**

**Tarih ve İmza**

## TEZ KONTROL ve BEYAN FORMU

30/06/2022

**Adı Soyadı:** Fisun KAŞKIR KESİN

**Anabilim Dalı:** Biyoistatistik Anabilim Dalı

**Tez Konusu:** Meta Analizinde Yanlılık Değerlendirme Yöntemlerinin Karşılaştırılması

<u>ÖZELLİKLER</u>	<u>UYGUNDUR</u>	<u>UYGUN DEĞİLDİR</u>	<u>AÇIKLAMA</u>
Tezin Boyutları	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Dış Kapak Sayfası	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
İç Kapak Sayfası	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Kabul Onay Sayfası	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Sayfa Düzeni	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
İçindekiler Sayfası	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Yazı Karakteri	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Satır Aralıkları	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Başlıklar	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Sayfa Numaraları	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Eklerin Yerleştirilmesi	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Tabloların Yerleştirilmesi	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Kaynaklar	<input checked="" type="checkbox"/>	<input type="checkbox"/>	

### DANIŞMAN ONAYI

**Unvanı Adı Soyadı:** Prof. Dr. İlker ERCAN

**İmza:**

## İÇİNDEKİLER

Dış Kapak	
İç Kapak	
ETİK BEYAN.....	II
KABUL ONAY.....	III
TEZ KONTROL ve BEYAN FORMU .....	IV
İÇİNDEKİLER .....	V
TÜRKÇE ÖZET .....	VII
İNGİLİZCE ÖZET .....	VIII
1. GİRİŞ .....	1
2. GENEL BİLGİLER.....	6
2.1. Meta Analizi.....	6
2.2. Meta Analizde Kullanılan Kavramlar .....	7
2.2.1. Ağırlık .....	7
2.2.2. Etki Büyüklüğü .....	7
2.2.3. İkili Verilerin Özet İstatistikleri .....	8
2.2.3.1. Relatif Risk (Risk Ratio, RR).....	8
2.2.3.2. Odds Oranı (Odds Ratio, OR).....	9
2.2.3.3. Risk Farkı (Risk Difference (RD)).....	10
2.2.3.4. Tedaviye Gerekli Sayı (NNT, The Number Needed To Treat) .....	11
2.3. Meta Analizde Heterojenlik .....	11
2.3.1. Cochran'ın Q İstatistiği ile Heterojenlik Testi .....	12
2.3.2. Heterojenlik Ölçümleri.....	14
2.3.2.1. H istatistiği .....	15
2.3.2.2. $\tau^2$ İstatistiği .....	16
2.3.2.3. R İstatistiği .....	18
2.3.2.4. $I^2$ İstatistiği.....	18
2.4. Meta Analizde Model Seçimi .....	20
2.4.1. Sabit Etki Modeli (Fixed Effect Model) .....	21
2.4.2. Rasgele Etkiler Modeli (Random Effects Model).....	22
2.5. İkili Verilerin Özet İstatistiklerinin Birleştirilmesinde Kullanılan Yöntemler ...	24
2.5.1. Mantel Haenszel Yöntemi.....	24
2.5.2. Peto Yöntemi .....	26
2.5.3. Genel Varyansa Dayalı Yöntem .....	27
2.5.4. DerSimonian-Laird Yöntemi .....	28
2.6. Meta Analizde Yanlılık.....	29
2.6.1. Meta Analizde Yanlılığa Neden Olabilen Yanlılık Çeşitleri .....	31
2.6.2. Yanlılıktan Kaçınmak İçin Yapılması Gerekenler .....	37
2.6.3. Literatürde Geliştirilen Meta Analizde Yanlılık Oluşturma Yaklaşımları....	40
2.6.4. Meta Analizde Yanlılık Belirlemede Kullanılan Yöntemler .....	41

2.6.4.1. Funnel Grafiđi .....	41
2.6.4.2. Funnel Grafiđi Asimetrisinin Deđerlendirmesinde Kullanılan Yöntemler ....	43
<b>3. GEREÇ VE YÖNTEM .....</b>	<b>50</b>
3.1. Simülasyon Senaryoları .....	51
<b>4. BULGULAR .....</b>	<b>52</b>
4.1. Meta Analizinde Yanlılık Olmadıđı Duruma Ait Bulgular .....	52
4.2. Meta Analizinde %10 Yanlılık Oluřturulduđu Duruma Ait Bulgular .....	75
4.3. Meta Analizinde %20 Yanlılık Oluřturulduđu Duruma Ait Bulgular .....	92
4.4. Meta Analizinde %40 Yanlılık Oluřturulduđu Duruma Ait Bulgular .....	110
4.5. Meta Analizinde %60 Yanlılık Oluřturulduđu Duruma Ait Bulgular .....	128
<b>5. TARTIřMA ve SONUÇ .....</b>	<b>154</b>
<b>6. KAYNAKLAR .....</b>	<b>161</b>
<b>7. SİMGELER VE KISALTMALAR .....</b>	<b>165</b>
<b>8. TEřEKKÜR .....</b>	<b>166</b>
<b>9. ÖZGEÇMİř .....</b>	<b>167</b>



## TÜRKÇE ÖZET

Meta analizi, aynı konu ile ilgili farklı araştırmacıların birbirinden farklı yer ve zamanlarda yapmış olduğu, farklı çalışma sonuçlarını uygun şekilde birleştiren bir yöntemdir. Meta analizi sonucu elde edilen bilgilerin yanlı olup olmadığının belirlenmesi için öznel bir değerlendirme sunan funnel grafiklerinin objektif olarak değerlendirilmesini sağlayan istatistiksel testler, meta analizi sonuçlarının güvenilirliğini değerlendirme imkânı sunar. Yanlılığı tespit etmeye yönelik istatistiksel testlerin performansının incelenmesi, ilgili istatistiksel testlerin kullanım zamanı, şekli ve uygulanmaya uygun oldukları araştırma koşulları hakkında bilgi sağlar.

Bu tez çalışmasında, ikili değer alan verilerin meta analizinde funnel grafiği asimetri değerlendirmesinde yanlılık belirlemek için kullanılan literatürde yer alan Begg, Egger, Thompson, Schwarzer ve Harbord testlerinin Tip-I hata oranları ve güçleri yönünden performansları incelenmiştir. Testlerin Tip-I hatayı koruma ve güçleri yönünden performanslarının farklı çalışma sayıları, farklı örneklem hacimleri, farklı yanlılık dereceleri ve farklı hastalık-sonuç oranları durumlarında nasıl etkilendiği simülasyon senaryoları ile incelenmiştir.

Yapılan simülasyon çalışmaları sonucunda Schwarzer ve Harbord testleri funnel grafiği asimetrisini tanımlamak için düşük istatistiksel güce, Begg, Egger ve Thompson testleri ise funnel grafiği asimetrisi olmadığında şişirilmiş Tip-I hata oranlarına sahiptir. Ele alınan koşullarda, testler arasında her tür veri için kullanılabilir yanlılığı saptamak için en iyi test yoktur. Testlerin performansları meta analizine dahil edilen çalışmaların sayısı, yanlılık seviyeleri, hastalık-sonuç oranları ve örneklem hacimleri ile değiştiğinden yanlılığı saptamaya yönelik yöntemler seçilirken bu değişkenlerin dikkate alınması gerekmektedir. Bu tez çalışmasında ele alınan simülasyon senaryoları sonucunda Tip-I hata olasılık değerini koruma yönünde en iyi performansı meta analizine alınan çalışma sayısı, örneklem hacmi az ise Schwarzer testi, fazla ise Harbord testinin gösterdiği söylenebilir. Simülasyon senaryoları sonucunda testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, en iyi performansı Egger testi göstermiştir.

**Anahtar kelimeler:** Meta Analizi, Funnel grafiği, Yanlılık, Funnel grafiği asimetri testleri



## İNGİLİZCE ÖZET

### COMPARISON OF BIAS EVALUATION METHODS IN META ANALYSIS

Meta analysis is a method that combines the results of different studies conducted by different researchers at different places and times on the same subject. Statistical tests that provide objective evaluation of funnel plots, which provide a subjective evaluation to determine whether the information obtained as a result of meta analysis is biased, offers the opportunity to evaluate the reliability of meta analysis results. Examining the performance of statistical tests to detect bias provides information about the time and manner of use of the relevant statistical tests, and the research conditions in which they are appropriate to be applied.

In this thesis, the performances of the Begg, Egger, Thompson, Schwarzer and Harbord tests in the literature, which are used to determine the bias in the funnel plot asymmetry evaluation in the meta analysis of binary data, were examined in terms of Type-I error rates and powers. How the performances of the tests in terms of Type-I error protection and powers were affected in cases of different number of studies, different sample sizes, different degrees of bias and different disease-outcome ratios were examined with simulation scenarios.

As a result of the simulation studies, Schwarzer and Harbord tests have low statistical power to define funnel plot asymmetry, and Begg, Egger and Thompson tests have inflated Type-I error rates when there is no funnel plot asymmetry. In the given conditions, there is no best test for detecting bias that can be used for all types of data between tests. Since the performance of tests varies with the number of studies included in the meta analysis, levels of bias, disease-outcome ratios, and sample size, these variables need to be taken into account when choosing methods for detecting bias. As a result of the simulation scenarios discussed in this thesis, it can be said that the Schwarzer test shows the best performance in terms of preserving the Type-I error probability value if the number of studies included in the meta-analysis and the sample size is low, and the Harbord test if it is more. As a result of the simulation scenarios, when the performance of the tests in terms of detecting bias was evaluated, the Egger test showed the best performance.

**Keywords:** Meta Analysis, Funnel plot, Bias, Funnel plot asymmetry tests

## 1. GİRİŞ

Meta analizi, sistematik bir deęerlendirmede yer alan alıřma sonularının istatistiksel olarak sentezlenmesidir. Bazen, bireysel bir alıřma sonu vermez ya da eliřkili sonular verebilir; aynı soruları soran birka alıřma yapılmıř ve genel bir sonuca ihtiya duyulmuř olabilir; arařtırmacılar zaman, maliyet ya da bir hastalıęın veya ilgili bir zellięin nadir grlmesi gibi rneklem hacimlerini kısıtlayan nedenlerden dolayı kk rneklem hacimleri ile alıřmak zorunda kalarak, alıřmaları iin gerekli ve yeterli gce ulařamayabilirler. Meta analizi, bu gibi durumlarda uygulanan, anaktle parametresine ynelik daha genel ve zet bir deęere ulařılmasını saęlamak iin sonuları tek bir zetlenmiř etki byklę elde etmek zere birleřtiren bir yaklařımdır. Meta analizinin sonuları, etki tahminlerinin kesinlięini arttırabilir, bireysel alıřmaların ortaya koyamadıęı soruları yanıtlayabilir, grnřte uyuřmayan alıřmalardan kaynaklanan tartıřmaları zebilir ve yeni hipotezler oluřturabilir (Haidich, 2010).

Meta analizinde, doęru sonular elde edebilmek iin ilk yapılması gereken, analize dahil edilecek alıřmaların sistemli ve dikkatli bir řekilde seilmesidir. Meta analizinde kullanılmak zere seilen alıřmaların oluřturduęu rneklem, anaktleyi temsil etmemesi nedeniyle oluřan yanlılıęın meta analizi sonularının gvenilirlięini olumsuz ynde etkileme riski vardır. Bu nedenle hangi alıřmaların meta analizine seileceęi nemli bir endiře kaynaęı oluřturur. Bir meta analizinde yanlılıęın arařtırılması iin birincil grsel ara asimetrisinin deęerlendirildięi funnel grafikleridir, ancak bir funnel grafięi farklı arařtırmacılar tarafından farklı řekilde yorumlanabilir. Arařtırmacılar iin meta analizinde yanlılık belirlemede znel bir deęerlendirme sunan funnel grafiklerini objektif olarak deęerlendirme ihtiyaı sonucunda eřitli istatistiksel testler geliřtirilmiřtir. Begg, Egger, Schwarzer, Harbord, Macaskill, Peters, Thompson, Arcsin-Begg, Arcsin-Egger, Arcsin-Thompson

yöntemleri ile geliştirilmiş istatistiksel testler, funnel grafiğinin asimetrisinin objektif olarak değerlendirmesine imkân sunan istatistiksel testlerden bazılarıdır.

Funnel grafiği ve ilgili istatistiksel testler sonucunda ciddi bir yanlılığın olduğu durumda, araştırmacılar tarafından ilk sonuçlar temkinli bir şekilde dikkate alınır. Böylece yararsız, hatta zararlı olabilecek bir müdahale veya politika önerme gibi potansiyel ciddi hatalardan ve boşa giden kaynaklardan kaçınılabılır. Sadece üzerinde çalışılan meta analizinin bütünlüğü için değil, aynı zamanda alanın bütünlüğünü sağlamak için de yanlılık değerlendirilmelidir. Yanlılık olasılığını göz ardı eden ve daha sonra yanlış olduğu tespit edilen bir meta analizi yayınlandığında, meta analizine olan güven azalır. Yanlılığının değerlendirilmesi ile meta analizinin kullanımı ve kullanışlılığı da artacaktır (Rothstein, Sutton, & Borenstein, 2005). Funnel grafiği ve ilgili istatistiksel testler, meta analizinde yanlılığı saptamak için yaygın olarak kullanılmasına rağmen bunların bazıları iyi bilinmemekte ve hem istatistiksel hem de klinik literatürde doğru şekilde kullanılmamaktadır. Bu nedenle, funnel grafiğinin asimetrisinin değerlendirilmesinde kullanılan bu testlerin sonuçlarının yorumlanması yanıltıcı olabilmektedir. Funnel grafiğinin yorumlanmasını kolaylaştırabilecek, meta analizinde funnel grafiğiyle ilgili istatistiksel testlerin ne zaman ve nasıl kullanılacağı hakkında bulunan öneriler ise sınırlıdır. Bu testlerin özelliklerini değerlendirmek ve karşılaştırmak üzere yapılan bazı çalışmalar aşağıda yer almaktadır.

Begg ve Mazumdar (1994), bir meta analizinde yanlılığı belirlemek üzere, etki tahminleri ile örnekleme varyansları arasındaki ilişkiyi incelemek için funnel grafiğini tamamlamaya yönelik düzeltilmiş bir sıra korelasyon yöntemi önererek, çalışma özelliklerini simülasyonlar yoluyla değerlendirmiştir. Meta analizine alınan çalışma sayısı, seçim mekanizmasının niteliği, etki büyüklüğü tahminlerinin varyans aralığı ve gerçek temel etki büyüklüğünün testin gücünün belirlenmesinde etkili olduğu gözlenmiştir. Yetmiş beş bileşen içeren bir çalışma ile büyük meta analizleri için sıra korelasyon testinin “oldukça güçlü” olduğu, ancak 25 bileşen içeren bir çalışma ile orta derecede güce sahip olduğu sonucuna varılmıştır.

Egger ve arkadaşları (1997), meta analizlerinde funnel grafiği asimetrisini test etmek için bağımlı değişken olarak etki büyüklüğünün standartlaştırılmış tahmini ve bağımsız değişken olarak ise duyarlılığı kullanılarak bir regresyon modeli

tanımlayarak, yayınlanmış meta analizlerinde yanlılığın yaygınlığını değerlendirmişlerdir. Önde gelen genel tıp dergilerinde yayınlanan meta analizlerin %38'inde ve Cochrane Sistemik İncelemeler Veri tabanındaki incelemelerin %13'ünde funnel grafiği asimetrisi bulunmuştur. Sonuç olarak, analiz, meta analizlerinde yanlılığın varlığı için yararlı bir test sağlamıştır.

Duval ve Tweedie (2000), meta analizinde bulunabilecek eksik çalışmaların sayısını ve bu çalışmaların sonuçlar üzerindeki etkisini tahmin etmek için parametrik olmayan bir yöntem olarak Trim-Fill yöntemini geliştirmiştir. Trim-Fill yöntemi, funnel grafiklerinin kullanımını resmileştiren, sıralamaya dayalı basit veri artırma tekniğidir. Eksik çalışmalar için düzeltme yapıldıktan sonra, genel etki büyüklüğünün nokta tahmininin yaklaşık olarak doğru olduğunu ve etki büyüklüğüne ait güven aralıklarının kapsamının büyük ölçüde iyileştirildiğini, çoğu durumda nominal güven sınırlarının tamamen düzeldiğini tespit etmişlerdir.

Sterne ve arkadaşları (2000), yaptıkları simülasyon çalışmalarında etki büyüklüğü ve standart hatası arasında doğrusal bir ilişki olduğunu varsayarak, sıra korelasyonunun ve regresyon yöntemlerinin özelliklerini karşılaştırmıştır. Yetmiş sekiz meta analizinden oluşan bir örnekleme dayanılarak, 5, 10, 20 ve 30 çalışma içeren "tipik" meta analizleri türetilmiş, orta şiddette yanlılık, ağır yanlılık ve yanlılığın olmadığı durumlarda ağırlıklı regresyon yönteminin ve sıra korelasyon testinin gücünü değerlendirmek için simülasyonlar yapılmıştır. Yazarlar, tıp alanında yapılan meta analizi çalışmalarında tipik olan durumlarda regresyon yönteminin, sıra korelasyon yönteminden daha güçlü olduğu sonucuna varmışlardır. Yanlılığı tespit etme gücü, meta analizine alınan çalışma sayısı arttıkça artmıştır. Bununla birlikte, regresyon yöntemi, büyük etki büyüklükleri, çalışma başına az sayıda olay ve bütün çalışmaların benzer büyüklükte olmaları gibi bazı durumlarda çok yüksek olan yanlış pozitif oranları vermiştir.

Macaskill ve arkadaşları (2001), yanlılık için Begg'in yöntemi, Egger'in yöntemi, örneklem hacmi ve etki büyüklüğünün regresyonu (funnel grafiği regresyonu) olmak üzere üç basit test yönteminin performansını simülasyon çalışması ile karşılaştırmışlardır. Testler, yanlılığın varlığında ve yokluğunda simüle edilmiş meta analizlerine uygulanmış, sonuç olarak, testlerin hiçbirinin tutarlı bir şekilde iyi

performans göstermediği raporlanmıştır. Test performansı, gerçek etki büyüklüğünün büyüklüğüne, çalışma büyüklüğünün dağılımına ve bir veya iki kuyruklu anlamlılık testi uygulanıp uygulanmadığına göre değişmiştir. Bununla birlikte, Tip-I hata oranları sonuçlarına dayanarak, birleştirilmiş varyansın tersine ağırlıklandırılan funnel grafiği regresyonu tercih edilen yaklaşım olmuştur.

Harbord ve arkadaşları (2006), Etkinlik skoru (efficient score) ve varyansına dayanarak funnel grafiği asimetrisi için değiştirilmiş bir doğrusal regresyon testi geliştirmişlerdir. Bu testin performansı, yayınlanan kontrollü çalışmaların özelliklerine dayanan simülasyonlarda önerilen diğer testlerle karşılaştırılmıştır. Denemeler arası heterojenlik az olduğunda veya hiç olmadığında, bu modifiye edilmiş test, Egger testine benzer gücü korurken nominal seviyeye yakın yanlış pozitif orana sahiptir. Denemeler arası heterojenliğin derecesi büyük olduğunda, önerilen testlerin hiçbirinin eşit derecede iyi özelliklere sahip olmadığı belirtilmiştir.

Peters ve arkadaşları (2006), Egger'in yönteminin performansını, örneklem hacmine (Macaskill'in testinin bir modifikasyonu) dayalı bir regresyon testiyle karşılaştırmışlardır. Yanlılığın ve çalışmalar arasındaki heterojenliğin varlığında ve yokluğunda bir dizi senaryo altında meta analizlerin simülasyonu yapılmıştır. Her regresyon testi için Tip-I hata oranları ve mevcut olduğu zaman yanlılığı saptama güçleri ölçülmüştür. Egger'in regresyon testi için Tip-I hata oranlarının, Peters'in regresyon testi için olanlardan daha yüksek çıktığı raporlanmıştır. Peters'in testinin, altta yatan OR değerinin büyüklüğüne, meta analizindeki birincil çalışmaların sayısına ve çalışmalar arası heterojenliğin seviyesine bakılmaksızın uygun Tip-I hata oranlarına sahip olduğu bildirilmiştir. Peters'in testinin, Egger'in regresyon testine benzer bir güce sahip olduğu ifade edilmiştir. Uygun Tip-I hata oranları ve lnOR ve varyansı arasındaki korelasyondaki azalma nedeniyle, özet tahminler lnOR'ler olduğunda Peters'in testinin, Egger regresyon testi yerine kullanılabileceği belirtilmiştir.

Sterne ve arkadaşları (2011), makalelerinde, randomize kontrollü çalışmaların meta analizlerinde funnel grafiği asimetrisinin nasıl inceleneceğini ve yorumlanması gerektiğini önermektedir. Öneriler, 2007 yılına kadar yayımlanan literatürün ayrıntılı bir MEDLINE incelemesine ve metodologlar arasındaki tartışmalara dayanmaktadır.

Jin ve arkadaşları (2014), istatistiksel ilkeler, uygulama ve yazılım dahil olmak üzere meta analizinde yanlılık ile başa çıkma yöntemleri ile bu yöntemlerin avantajları ve sınırlamaları hakkında kritik ve kapsamlı bir tartışma sağlayarak, doğum sırasında kadınlara sürekli desteğin meta analizinde, yöntemlerin pratik bir uygulamasını göstermişlerdir. Güvenilir bir test kullanılarak meta analizinde yanlılıktan şüphelenildiğinde, duyarlılık analizi yapmak için Copas seçim modelinin kullanılması önerilmiştir.

Kale ve Nirpharake (2017), yanlılığı tespit etmenin 4 yöntemini (Begg'in yöntemi, Egger'in yöntemi, Fail Safe N, Trim-Fill) benzetilmiş veriler kullanılarak karşılaştırmıştır. Sonuçlar, Fail Safe N ve Trim-Fill yöntemlerinin, meta analizine dahil edilen çalışmaların sayısı az olduğunda daha iyi performans gösterdiğini ortaya koymuştur. Öte yandan, meta analizine dahil edilen çalışmaların sayısı çok fazla olduğunda, Begg'in ve Egger'in yöntemlerinin daha iyi sonuçlar verdiği belirtilmiştir.

Debray ve arkadaşları (2017), funnel grafiği asimetri testlerinin uygulanmasının (Egger'in yöntemi, Macaskill'in yöntemi, Peters'in yöntemi), survival verilerinin meta analizinde daha az uygun olabileceğini ifade etmiştir. Daha sonra, toplam gözlemlenen olay sayısına dayanan ve çarpımsal bir varyans bileşeni benimseyen iki yeni test önermişlerdir. Sonuçlar, funnel grafiği asimetrisini saptamak için önceden bilinen testlerin, özellikle çalışmalar, katılımcının çıkışından etkilendiğinde, survival verilerinin meta analizinde düşük güç veya aşırı Tip-I hata oranlarından sıkıntılı olduğunu göstermiştir. Debray ve arkadaşlarının (2017) önerdiği yeni test makul bir güç sağladığından ve uygun Tip-I hata oranlarını koruduğundan, survival verilerinin meta analizinde funnel grafiği asimetrisini değerlendirmek için kullanılmasını tavsiye etmişlerdir.

Bu tez çalışmasında, ikili değer alan verilerin meta analizinde funnel grafiği asimetri değerlendirmesi için kullanılan parametrik ve parametrik olmayan yanlılık belirleme testlerinin Tip-I hata oranları ve güçleri yönünden performanslarının farklı çalışma sayıları, farklı örneklem hacimleri, farklı yanlılık dereceleri ve farklı hastalık-sonuç oranları durumlarında nasıl etkilendiğinin simülasyon çalışması ile değerlendirilmesi amaçlanmıştır. Çalışmaya Begg'in, Egger'in, Schwarzer'in, Harbord'in, Thompson'in meta analizinde yanlılık belirleme testleri dahil edilmiştir.

## 2. GENEL BİLGİLER

### 2.1. Meta Analizi

Literatürde aynı konu ile ilgili farklı araştırmacıların birbirinden farklı sonuçlar gösteren yayınları olduğu görülebilmektedir. Meta analizi, farklı çalışma sonuçlarına göre bir sentez yapabilmek için istatistiksel yöntem gereksinim duyulmasından dolayı ortaya çıkmış ve geliştirilmiştir. Meta analizi diğer derleme yöntemlerinden farklı olarak, araştırma sonuçlarını nicel olarak sunmaktadır.

Meta analizi belirli bir konu ile ilgili toplam kanıtları özetlemek için önceki araştırma sonuçlarını istatistiksel olarak birleştiren ve eleştirel olarak değerlendiren bir yöntemdir. Meta analizi yerine eş anlamlı olarak genel değerlendirme terimi de kullanılmaktadır (Spector, & Thompson, 1991). Farklı yer ve zamanda değişik sayıda birim üzerinde yapılan araştırmaların sonuçları uygun biçimde birleştirilirse değişkenin anakütledeki dağılımı da uygun olarak belirlenebilmektedir. Sonuç olarak, geçerli, tutarlı ve minimum varyanslı tahminler yapılarak anakütleyle ilişkin doğru kararlar alınabilmektedir (Hunter, & Schmidt, 2004).

Meta analizinin başlıca amaçları aşağıdaki gibi sıralanabilir (Haidich, 2010; Normand, 1999);

- i. Literatürde ortaya çıkan çelişkileri incelemek ve nedenlerini bulmak,
- ii. Küçük hacimli örneklerle icra edilmiş çalışmaları birleştirerek genel örneklem hacmini artırıp parametre tahminlerinin gücünü ve kesinliğini artırmak,
- iii. Yapılan farklı araştırmalar arasında gözlenen heterojenliğin doğru kaynaklarını tespit etmek,
- iv. Bireysel araştırmalarda göz önüne alınmayan ancak etkili olabileceği düşünülen değişik alt gruplardaki işlem etkinliğini araştırmak,

v. Sonradan yapılacak olan arařtırmalar için bilgi kaynađı olmak,

vi. Ulařılan sonuçlarla ileriki arařtırmalara ışık tutmak ve yeni konuları arařtırmaya olanak sađlamak.

Meta analizinde yer alan adımlar (Kale, & Nirpharake, 2017):

- Her alıřma için etki büyüklüğü hesaplama
- Birleřik etki büyüklüğü hesaplama
- Heterojenlik testi
- Yanlılık testi
- Duyarlılık analizidir.

## **2.2. Meta Analizinde Kullanılan Kavramlar**

### **2.2.1. Ađırlık**

Genellikle, meta analizine alınan her bir alıřma, farklı örneklem hacimlerinde olmaktadır. Büyük örneklem hacimlerindeki alıřmaların tahminlerinin, daha küçük örneklem hacimlerindeki alıřmaların tahminlerinden daha iyi olma olasılığı yüksektir. Analize alınan alıřmalardaki toplam birim sayısının, meta analizine esas alınan tüm alıřmalardaki toplam birim sayısına oranlanmasıyla her bir alıřmanın meta analizindeki ađırlığı elde edilir.

### **2.2.2. Etki Büyüklüğü**

Literatürde etki büyüklüğü veya etki genişliği olarak karřımıza ıkabilen kavram, iki deđiřken arasındaki iliřki miktarı veya iki grup arasındaki farklılığın miktarı olarak kullanılır. Meta analizine dahil edilen her bir alıřma sayısal olarak birleřtirilmeden önce alıřmaların her biri için sonuç deđiřkenine ait ölçümlerin (etki büyüklüklerinin) hesaplanması gerekmektedir (Deeks, 2002). Etki büyüklükleri, farklı alıřmalardan elde edilen özet istatistiklerin ortak bir ölçüye dönüřtürülmesini sađlamasından dolayı önemlidir.

Literatürde verinin biçimine göre deđiřiklik gösteren birçok etki büyüklüğü türü vardır. Meta analizine dahil edilen her bir alıřmanın sonuç deđiřkeni ikili deđer alan deđiřken ise hasta (ya da tedavi) ve kontrol grubunda ilgilenilen durumun görülme



sıklıkları arasındaki fark değerlendirilmektedir. Eğer değişken sürekli ise, iki bağımsız değişken arasındaki ilişkinin miktarı veya grup ortalamaları arasındaki farklılığın miktarı değerlendirilmektedir (Deeks, 2002).

### 2.2.3. İkili Verilerin Özet İstatistikleri

İkili veriler (dichotomous veya binary), her bireyin sonucunun sadece iki kategorik olası yanıtta biri olduğu verilerdir. İkili sonuç verileri, her katılımcı için sonuç ölü-canlı veya klinik iyileşme var-klinik iyileşme yok gibi iki olasılıktan biri olduğunda ortaya çıkar. İkili verilerle klinik çalışmalarda en sık kullanılan özet istatistikleri relatif risk (RR), odds oranı (OR), risk farkı (RD) ve tedaviye gerekli olan sayı (NNT) olarak bilinmektedir. İkili veri için tanımlayıcı istatistiklerin hesaplanmasında Tablo 1’deki kontenjans tablosu kullanılmaktadır.

Tablo 1. İkili veriler için etki büyüklüklerinin hesaplanmasında kullanılan kontenjans tablosu

		Sonuç		Toplam
		+	-	
Etken	Hasta(Tedavi) (H+)	a	b	(a+b=e)
	Kontrol (H-)	c	d	(c+d=f)
Toplam		(a+c=g)	(b+d=h)	n (Toplam)

Tablo 1’de verilen simgeler, hasta grubundaki durum riski olasılıklarını hesaplamak için eşitlik (2.1), kontrol grubundaki durum riski olasılıklarını hesaplamak için eşitlik (2.2)’de olduğu gibi kullanılabilir.

$$P_H = \frac{a}{(a + b)} \quad (2.1)$$

$$P_K = \frac{c}{(c + d)} \quad (2.2)$$

#### 2.2.3.1. Relatif Risk (Risk Ratio, RR)

Hasta grubundaki ilgilenilen durumun görülme sıklığının, kontrol grubunda ilgilenilen durumun görülme sıklığına oranlanması ile hesaplanır. Diğer bir ifadeyle, hasta grubu ile kontrol grubunun risklerinin oranıdır ve risk oranı olarak da ifade edilir. Tablo 1’den yararlanılarak eşitlik (2.3)’de verilen formül ile RR elde edilebilir.

$$RR = \frac{\left(\frac{a}{(a+b)}\right)}{\left(\frac{c}{(c+d)}\right)} \quad (2.3)$$

Hasta ve kontrol gruplarındaki ilgilenilen durumun gözlenme sıklıkları birbirine eşit ise RR değeri 1'e eşit olur. Hasta grubunda ilgilenilen durumun gözlenme sıklığı, kontrol grubundakinden büyükse RR değeri 1'den büyük olur ve risk artışı anlamına gelir. Benzer şekilde hasta grubunda ilgilenilen durumun gözlenme sıklığı, kontrol grubundakinden küçükse RR değeri 1'den küçük çıkacağından riskin azaldığı söylenebilmektedir (Süt, & Şenocak, 2007). Güven aralığı risk oranına göre simetrik değildir. RR negatif değerler almazken, oldukça büyük değerler alabilir.

Logaritmik RR'nin güven aralığı yaklaşık olarak simetriktir. Bundan dolayı RR'nin güven aralığı logaritmik değerine göre hesaplanır.  $\ln(RR)$  için %95 güven aralığı,  $\ln(RR)$ 'ye ait varyans değeri ve RR için %95 güven aralığı sırasıyla eşitlik (2.4), eşitlik (2.5) ve eşitlik (2.6)'da verildiği gibi hesaplanır. Güven aralığı 1 değerini içermiyorsa RR değeri istatistiksel olarak anlamlıdır.

$$\%95 \text{ G. A.} = \ln(RR) \pm 1,96\sqrt{\text{Var}(\ln(RR))} \quad (2.4)$$

$$\text{Var}(\ln(RR)) = \frac{1}{a} - \frac{1}{(a+b)} + \frac{1}{c} - \frac{1}{(c+d)} \quad (2.5)$$

$$\%95 \text{ G. A.} = \exp(\ln(RR) \pm 1,96\sqrt{\text{Var}(\ln(RR))}) \quad (2.6)$$

### 2.2.3.2. Odds Oranı (Odds Ratio, OR)

Odds oranı, hasta grubunda ilgilenilen durumun gerçekleşme odds'unun kontrol grubundaki odds'a oranlanmasıyla elde edilir. Diğer bir ifadeyle, bir olayın olma olasılığının, olmama olasılığına oranı odds oranıdır. Odds oranı, eşitlik (2.7)'de olduğu gibi hesaplanmaktadır.

$$OR = \frac{\frac{a}{(a+b)}}{\frac{c}{(c+d)}} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a \cdot d}{b \cdot c} \quad (2.7)$$

RR'de olduğu gibi OR pozitif değerler alabilirken, negatif değerler alamadığından, OR'nin de simetrik bir güven aralığı yoktur. Bu nedenle RR değerinde olduğu gibi OR değerinin güven aralığı hesabında da logaritmik değeri kullanılır. OR'nin %95 güven aralığı ve  $\ln(OR)$ 'ye ait varyans değeri sırasıyla eşitlik (2.8) ve eşitlik (2.9)'daki gibi hesaplanır (Borenstein, Hedges, Higgins, & Rothstein, 2009).

$$\%95 \text{ G. A.} = \exp(\ln(OR) \pm 1,96\sqrt{\text{Var}(\ln(OR))}) \quad (2.8)$$

$$\text{Var}(\ln(OR)) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} \quad (2.9)$$

OR değerinin 1 değerinden küçük olması, etkenin hastalık riskini azalttığı, 1 değerine eşit olması risk faktörünün hastalık üzerinde etkisinin olmadığı, 1 değerinden büyük olması ise hastalık riskini artırdığı anlamına gelmektedir (Süt, & Şenocak, 2007). Güven aralığı 1 değerini içermiyorsa OR değeri istatistiksel olarak anlamlıdır.

### 2.2.3.3. Risk Farkı (Risk Difference (RD))

Hasta ve kontrol grubundan oluşan ve sonucu ikili değer alan yapıda olan bir çalışmada, etkene maruziyet durumunun kontrol grubuna göre etkinliğini tespit etmenin en basit yolu, hasta ve kontrol gruplarındaki ilgilenilen durumun gerçekleşme riskleri arasındaki farkları hesaplamaktır. Risk farkı, mutlak risk farkı, atfedilen görel risk (attributable relative risk) ve atfedilen risk olarak da adlandırılmaktadır. Risk farkı, kontrol grubundaki ilgilenilen durumun gerçekleşme riskinden hasta grubundaki riskin çıkarılmasıyla hesaplanmaktadır (Süt, & Şenocak, 2007).

$$RD = \frac{c}{(c+d)} - \frac{a}{(a+b)} \text{ veya } RD = \left| \frac{a}{(a+b)} - \frac{c}{(c+d)} \right| \quad (2.11)$$

Risk farkı pozitif ve negatif değerler alabilmektedir. Bu nedenle, logaritmik dönüşüme ihtiyaç yoktur. Risk farkı için %95 güven aralığı ve risk farkına ait varyans değeri sırasıyla eşitlik (2.12) ve (2.13)'de verildiği gibi hesaplanır. %95 güven aralığı 0'ı içermiyorsa RD istatistiksel olarak anlamlı olur (Süt, & Şenocak, 2007).

$$\%95 \text{ G. A.} = RD \pm 1,96\sqrt{\text{Var}(RD)} \quad (2.12)$$

$$Var(RD) = \frac{\left(\frac{a}{(a+b)}\right)\left(1 - \left(\frac{a}{(a+b)}\right)\right)}{(a+b)} + \frac{\left(\frac{c}{(c+d)}\right)\left(1 - \left(\frac{c}{(c+d)}\right)\right)}{(c+d)} \quad (2.13)$$

#### 2.2.3.4. Tedaviye Gerekli Sayı (NNT, The Number Needed To Treat)

Belli bir etkenin, sonucu oluşturma yönündeki etkisini (daha çok ilaçların tedavi etkisini), nesnel olarak görebilmek için, en az kaç olgunun ele alınması gerektiğini hesaplayan NNT formülü eşitlik (2.14)'de verilmiştir (Süt, & Şenocak, 2007).

$$NNT = \frac{1}{RD} = \frac{1}{\frac{c}{(c+d)} - \frac{a}{(a+b)}} \quad (2.14)$$

NNT değeri yalnızca pozitif değer almasına rağmen negatif değer gözleniyorsa bu duruma “yan etkiye gerekli sayı” (NNH, number needed to harm) denmekte ve NNH değeri tedavi yönteminin zararlı etkiye sahip olduğunu göstermektedir. NNT için %95 güven aralığı, RD'nin güven aralığının tersinin alınması ile hesaplanır (Süt, & Şenocak, 2007). NNT meta analizinde tercih edilen bir yöntem değildir.

### 2.3. Meta Analizde Heterojenlik

Meta analizi çalışmalarında heterojenlik söz konusu ise değişkenliğin iki kaynağı vardır. Bunlardan biri örnekleme hatası (çalışma içi değişkenlik), diğeri ise çalışmalar arasındaki değişkenliktir. Meta analizine dahil edilen her bir çalışma farklı örneklemlerden oluştuğu için örnekleme hatası yani çalışma içi değişkenlik meta analizinde daima mevcuttur (Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006). Dolayısıyla heterojenlik, çalışmalar arası varyans anlamı derecede arttığında ortaya çıkar.

Heterojenlik genelde, klinik, metodolojik ve istatistiksel olmak üzere üç kategoriye ayrılır. Klinik heterojenlik, deneklere yapılan müdahaleye, deneklerin özelliklerine ve çalışmadaki sonuç değişkeninin çeşitliliğine göre çalışmalardaki klinik farklılık olarak tanımlanmaktadır. Metodolojik heterojenlik, deneme düzeni ile

ilgili heterojenliktir. Çalışmanın kalitesi, çalışmanın süresi ve kullanılacak olan istatistiksel yöntemlerin farklılığı metodolojik heterojenliğin bir göstergesidir. İstatistiksel heterojenlik ise etki büyüklüklerindeki değişkenlikle ilişkilidir (Higgins, & Green, 2006).

Meta analizinde heterojenlik olduğu durumda, sonuçların yorumlanması ve değerlendirilmesi aşamalarının nasıl yapılacağı ile ilgili tam olarak bir fikir birliği bulunmamakla birlikte, çalışmalar homojen ise sabit etki modeli kullanılması, heterojenlik varsa, asla görmezden gelinmemesi önerilmektedir. Heterojenliğin olduğu durumda, uç çalışmaların tespit edilmesi gerekmekte ve analizler uç çalışmaların olup olmasına göre değerlendirilmelidir. Analiz aşamasında, çalışmalar arasında daha fazla homojenlik sağlamak ve heterojenliği ortadan kaldırmak amacıyla uç çalışmaların analiz dışında tutulması tavsiye edilmektedir. Uç çalışmaların analiz dışında bırakılması durumunda, bazı analizlerde homojenlik sağlanırken, bazılarında sağlanmayabilir. Bu durumda heterojenliğin kaynağı tespit edilemiyorsa rasgele etkiler modeli kullanılması önerilmektedir. Çalışmalardaki bireylerin karakteristik özelliklerinde, çalışmaya dâhil etme kriterlerindeki olası farklılığı tanımlamak için çalışmalar farklı homojen gruplar içerisinde sınıflandırılarak, öncelikle her bir homojen olduğu düşünülen alt grup içinde heterojenlik testi yapılmalıdır. Alt gruplarda homojenlik sağlanabilmişse sabit etki modeli, sağlanamamışsa da rasgele etkiler modeli kullanılmalıdır (Whitehead, 2002).

### 2.3.1. Cochran'ın Q İstatistiği ile Heterojenlik Testi

Heterojenliği test etmede kullanılan geleneksel ve kabul görmüş test Cochran'ın Q testidir.  $H_0: \theta_1 = \theta_2 = \dots = \theta_K = \theta$  (tüm çalışmaların etki büyüklükleri eşittir) ya da  $H_0: \sigma_{\text{ç.A.}}^2 = 0$  (çalışmalar arası varyans sıfıra eşittir) şeklinde kurulan yokluk hipotezi, “en az bir etki büyüklüğü farklıdır” veya “varyans sıfırdan farklılık göstermektedir” şeklinde olan hipotezlere karşı sınırdır. Q, (k-1) serbestlik derecesi ile ki-kare dağılımına sahiptir. Q test istatistiği eşitlik (2.15)'teki gibi ifade edilir.

$$Q = \sum_{i=1}^k w_i (Y_i - M)^2 \quad (2.15)$$

M, gözlenen etki büyüklüklerinin ağırlıklı ortalamasıdır ve eşitlik (2.15)'te yerine konulduğunda eşitlik (2.16) ve gerekli düzenlemeler yapıldığında eşitlik (2.17) elde edilir.

$$Q = \sum_{i=1}^k w_i \left( Y_i - \frac{\sum_{i=1}^k w_i Y_i}{\sum_{i=1}^k w_i} \right)^2 \quad (2.16)$$

$$Q = \sum_{i=1}^k w_i Y_i^2 - \frac{(\sum_{i=1}^k w_i Y_i)^2}{\sum_{i=1}^k w_i} \quad (2.17)$$

Elde edilen Q değeri,  $\chi_{k-1}^2$  dağılımına sahip değer ile karşılaştırılarak test edilir (Hardy, & Thompson, 1998; Rucker, Schwarzer, Carpenter, & Schumacher, 2008). Q test istatistiğinin beklenen değeri eşitlik (2.18)'deki gibi hesaplanır.

$$E[Q] = \sigma_{\zeta.A.}^2 \left( \sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i} \right) + (k - 1) \quad (2.18)$$

Q ile  $\hat{\sigma}_{\zeta.A.}^2$  (çalışmalar arası varyansın tahmini değeri) eşitlik (2.19)'da verilen şekilde hesaplanabilir (Hardy, & Thompson, 1998; Mittlböck, & Heinzl, 2006).

$$\hat{\sigma}_{\zeta.A.}^2 = \frac{Q - (k - 1)}{\left( \sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i} \right)} \quad (2.19)$$

Eğer  $Q < (k - 1)$  ise  $\hat{\sigma}_{\zeta.A.}^2$  negatif bir değer alır. Böyle bir durumda  $\hat{\sigma}_{\zeta.A.}^2$  değeri  $\max(0, \hat{\sigma}_{\zeta.A.}^2)$  şeklinde kullanılır. Böylece, rasgele etkiler modelinin etki büyüklüğünün ağırlığı ( $\frac{1}{V_R}$ ), sabit etki modelinin etki büyüklüğünün ağırlığından ( $\frac{1}{V_F}$ ) daha büyük olmayacaktır (Mittlböck, & Heinzl, 2006).

Çalışma içi varyansın tahmini iki farklı tahminle elde edilmektedir. İlk tahmin Takkouche ve arkadaşları (1999) tarafından eşitlik (2.20)'deki gibi hesaplanmıştır.

$$\hat{\sigma}_{\zeta.i.1}^2 = kv_F = \frac{k}{(\sum_{i=1}^k w_i)} \quad (2.20)$$

İkinci tahmin ise eşitlik (2.21)'deki gibi hesaplanır.

$$\hat{\sigma}_{\zeta.i.2}^2 = \frac{(k-1) \sum_{i=1}^k w_i}{\left(\sum_{i=1}^k w_i\right)^2 - \sum_{i=1}^k w_i^2} \quad (2.21)$$

Çalışma içi varyans tahminine ( $\sigma_{\zeta.i.2}^2$ ), dayalı Q test istatistiğinin beklenen değeri eşitlik (2.22)'deki gibi hesaplanır (Mittlböck, & Heinzl, 2006).

$$E[Q] = (k-1) \left( \left( \frac{\sigma_{\zeta.A.}^2}{\sigma_{\zeta.i.2}^2} \right) + 1 \right) \quad (2.22)$$

( $\sigma_{\zeta.i.2}^2$ ) değeri hem heterojenlik testinin gücü hem de  $E[Q]$  ile doğrudan ilişkilidir. Yani, ( $\sigma_{\zeta.i.2}^2$ )'nin kullanımı heterojenlik etkisini daha iyi göstermektedir. Dolayısıyla, ( $\sigma_{\zeta.i.2}^2$ ), heterojenlik testinin gücü üzerine çalışma içi varyansın etkisini değerlendirmek için ( $\sigma_{\zeta.i.1}^2$ )'e göre daha çok tercih edilir. Eğer  $w_1 = w_2 = \dots = w_k$  ise  $\sigma_{\zeta.i.1}^2 = \sigma_{\zeta.i.2}^2$  olduğuna dikkat etmek gerekir (Mittlböck, & Heinzl, 2006).

Literatürde, Cochran Q istatistiğinin anlamlılık seviyesi için 0,05 yerine 0,10 alınması önerilmektedir (Petitti, 2001). Meta analizine alınan çalışma sayısı az ise ( $k < 20$ ) Cochran Q testinin gücü düşük, fazla ise testin gücü yüksek olmaktadır. Ayrıca, çalışma sayısı arttıkça önemsiz olan küçük miktardaki heterojenlikte bile Q değeri de artar ve bu değer test edildiğinde istatistiksel olarak anlamlı kabul edilebilir. Bu durumda heterojenlik olup olmadığına karar vermek zorlaşacaktır. Bu yüzden, heterojenlik olduğunda, heterojenliğin miktarının belirlenmesi için Q test istatistiğinin yanında, heterojenlik ölçüsü de kullanılmalıdır (Mittlböck, & Heinzl, 2006; Rucker ve ark., 2008).

### 2.3.2. Heterojenlik Ölçümleri

Heterojenliğin varlığında, miktarının belirlenmesi için heterojenlik ölçümleri mevcuttur. Bu ölçümlerden en çok kullanılan,  $H^2$ ,  $R^2$ ,  $\tau^2$  ve  $I^2$  istatistikleridir. Bu istatistiklerden  $H^2$  ve  $R^2$  istatistiklerinin karekökü kullanılır (Mittlböck, & Heinzl, 2006; Rucker ve ark., 2008).

### 2.3.2.1. H İstatistiği

Higgins ve Thompson (2002) tarafından ortaya atılan H istatistiği Cochran Q istatistiği yardımıyla eşitlik (2.23) gibi verilmektedir (Higgins, & Thompson, 2002).

$$H^2 = \begin{cases} \frac{Q}{(k-1)}, & Q > (k-1) \\ 1, & Q \leq (k-1) \end{cases} \quad (2.23)$$

$H^2$  değeri 1 ile  $\infty$  arasındadır ve H değerinin 1 olması homojenliğin mükemmel olduğu anlamına gelmektedir. H değeri, çalışma sayısı ile sistematik olarak artış gösterir. H istatistiğinin beklenen değeri ve %95 güven aralığı sırasıyla eşitlik (2.24) ve (2.25)'teki gibi verilmiştir (Mittlböck, & Heinzl, 2006; Rucker ve ark., 2008).

$$E[H^2] = \frac{E[Q]}{(k-1)} = \left( \left( \frac{\sigma_{\text{Ç.A.}}^2}{\sigma_{\text{Ç.i.2}}^2} \right) + 1 \right) = \left( \frac{\sigma_{\text{Ç.A.}}^2 + \sigma_{\text{Ç.i.2}}^2}{\sigma_{\text{Ç.i.2}}^2} \right) \quad (2.24)$$

$$\exp[\ln(H) \pm 1,96 SE(\ln(H))] \quad (2.25)$$

Eşitlik (2.25)'teki  $SE(\ln(H))$ , Q istatistiğinin çalışma sayısından büyük ya da küçük olma durumuna göre farklı şekilde hesaplanmaktadır.  $Q > k$  olduğu durumda eşitlik (2.26),  $Q \leq k$  olduğu durumda eşitlik (2.27)'deki gibi hesaplanır (Higgins, & Thompson, 2002).

$$SE(\ln(H)) = \frac{1}{2} \left( \frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{2k-3}} \right) \quad (2.26)$$

$$SE(\ln(H)) = \sqrt{\frac{1}{2(k-2)} \left( 1 - \frac{1}{3(k-2)^2} \right)} \quad (2.27)$$

H istatistiğinin değeri çalışma sayısı küçük olduğu durumlarda ( $k < 8$ ) küçük bir yanlılık gösterebilir. H istatistiğinin değişkenliği çalışma sayısı az olduğunda büyüktür ve heterojenliği ayırt etmek zorlaşır. Çalışma sayısı artırıldığında ise değişkenlik azalır. Bununla birlikte, homojenlikten bahsedebilmek için H istatistiğinin güven aralığı mutlaka 1 değerini içermelidir. Aksi durumda heterojenliğin varlığından bahsedilir (Higgins, & Thompson, 2002).

Uygulamada  $\max(1, H^2)$  olarak kullanılmaktadır.



$H^2$  istatistiğinin düzeltilmiş hali ise eşitlik (2.28)'de verilmiştir (Mittlböck, & Heinzl, 2006).

$$H_M^2 = \begin{cases} \frac{Q - (k - 1)}{(k - 1)}, & Q > (k - 1) \\ 0, & Q \leq (k - 1) \end{cases} \quad (2.28)$$

Formül homojenliğin sağlandığı durumda “0” değerini almaktadır. Uygulamada  $\max(0, H_M^2)$  olarak kullanılmakta ve 0 ile  $\infty$  arasında bir değer almaktadır. Düzeltilmiş  $H^2$  istatistiğinin beklenen değeri eşitlik (2.29)'da verilmiştir (Mittlböck, & Heinzl, 2006).

$$E[H_M^2] = \frac{\sigma_{C.A.}^2}{\sigma_{C.I.2}^2} \quad (2.29)$$

### 2.3.2.2. $\tau^2$ İstatistiği

$\tau^2$  istatistiği çalışmalar arası varyansı ifade etmektedir ve meta analizindeki rasgele etkiler modelinde olduğu gibi tahmin edilir. Meta analizinde  $\tau^2$  istatistiğinin tahmini için eşitlik (2.30)'de verilen DerSimonian Laird yönteminden yararlanılır (Borenstein ve ark., 2009).

$$\tau^2 = \frac{Q - (k - 1)}{C} \quad (2.30)$$

$$C = \sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i} \quad (2.31)$$

Dolayısıyla eşitlik (2.31) yardımıyla yeni  $\tau^2$  çalışmalar arası varyans eşitlik (2.32)'deki gibi elde edilir.

$$\hat{\tau}^2 = \begin{cases} \frac{Q - (k - 1)}{\sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i}}, & Q > (k - 1) \\ 0, & Q \leq (k - 1) \end{cases} \quad (2.32)$$

$\tau^2$  istatistiğinin dağılımı normal dağılıma uygunluk göstermemektedir.  $\tau^2$  istatistiğinin %95 güven aralığını elde edebilmek için literatürde bulunan çok sayıdaki yöntemden en basit olanına ait %95 güven aralığının alt ve üst sınır değerleri eşitlik (2.33) ve (2.34)'de sunulmuştur.

$$A = \exp\left(0,5 \ln\left(\frac{Q}{k}\right) - 1,96B\right) \quad (2.33)$$

$$U = \exp\left(0,5 \ln\left(\frac{Q}{k}\right) + 1,96B\right) \quad (2.34)$$

Alt ve üst sınır değerlerinde yer alan B standart hata değeri olup, çalışma sayısının büyük veya küçük oluşuna bağlı olarak Cochran Q istatistiğinden elde edilen farklı formüllerle hesaplanmaktadır. Eğer  $Q > k$  ise eşitlik (2.35),  $Q \leq k$  ise eşitlik (2.36)'da verildiği gibi hesaplanmaktadır.

$$B = \frac{1}{2} \left( \frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{2(k-3)}} \right) \quad (2.35)$$

$$B = \sqrt{\frac{1}{2(k-2) \left(1 - \frac{1}{3(k-2)^2}\right)}} \quad (2.36)$$

%95 Güven aralığına ait alt ve üst sınır değerleri formülleri eşitlik (2.37) ve eşitlik (2.38)'deki gibidir (Borenstein ve ark., 2009).

$$A_{\tau^2} = \frac{(k-1)(A^2 - 1)}{C} \quad (2.37)$$

$$U_{\tau^2} = \frac{(k-1)(U^2 - 1)}{C} \quad (2.38)$$

$\tau^2$  istatistiğinin, alt veya üst sınır değerlerinden herhangi biri "0"dan küçükse gerçek değeri yerine "0" değeri atanır. Eğer alt sınır değeri 0'dan büyük ise  $\tau^2$  istatistiksel olarak anlamlı kabul edilir (Borenstein ve ark., 2009).

### 2.3.2.3. R İstatistiği

R istatistiği daha çok rasgele etkiler modeline ait güven aralıklarının etkisini tanımlamakta kullanılır. Meta analizine dahil edilecek çalışma sayısından bağımsız olan yöntem H istatistiğine benzer,  $\tau^2$  istatistiğinden hesaplanır ve (2.39) eşitliğindeki gibi formüle edilir (Higgins, & Thompson, 2002; Rucker ve ark., 2008).

$$R^2 = \frac{\tau^2 + \sigma^2}{\sigma^2} \quad (2.39)$$

R değerinin “1” olması homojenliğin mükemmel olduğunu göstermektedir. R istatistiğinin %95 güven aralığı eşitlik (2.40)’daki gibi hesaplanmaktadır.

$$\exp[\ln(R) \pm 1,96 SE(\ln(R))] \quad (2.40)$$

R istatistiğinin ve %95 güven aralıklarının hesaplanmaları ve yorumlanmaları H istatistiğine benzerdir. Eşitlik (2.40)’daki  $SE(\ln(R))$ , Cochran Q istatistiğinin çalışma sayısından büyük ya da küçük oluşuna göre farklı şekilde hesaplanmaktadır. Buna göre,  $Q > k$  olduğu durumda eşitlik (2.41),  $Q \leq k$  olduğu durumda eşitlik (2.42)’de verilen şekilde hesaplanmaktadır (Higgins, & Thompson, 2002).

$$SE(\ln(R)) = \frac{1}{2} \left( \frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{2(k-3)}} \right) \quad (2.41)$$

$$SE(\ln(R)) = \sqrt{\frac{1}{2(k-2)} \left( 1 - \frac{1}{3(k-2)^2} \right)} \quad (2.42)$$

### 2.3.2.4. $I^2$ İstatistiği

Higgins ve Thompson tarafından (2002) önerilen yöntem, Cochran Q ve  $H^2$  istatistiğinden yararlanarak oluşturulmuş bir istatistik olup eşitlik (2.43), (2.44) ve (2.45)’de olduğu gibi farklı şekillerde elde edilmektedir (Rucker ve ark., 2008).

$$I^2 = \begin{cases} \frac{Q - (k-1)}{Q} \cdot 100, & Q > (k-1) \\ 0, & Q \leq (k-1) \end{cases} \quad (2.43)$$

$$I^2 = \begin{cases} \frac{H^2 - 1}{H^2} \cdot 100, & Q > (k - 1) \\ 0, & Q \leq (k - 1) \end{cases} \quad (2.44)$$

$$I^2 = \begin{cases} \frac{c\tau^2}{Q} \cdot 100, & Q > (k - 1) \\ 0, & Q \leq (k - 1) \end{cases} \quad (2.45)$$

$I^2$  istatistiğinin beklenen değeri,

$$E(I^2) = \frac{\tau^2}{\tau^2 + \sigma^2} \quad (2.46)$$

şeklinde hesaplanmaktadır.  $\tau^2=0$  için  $Q$  yaklaşık olarak  $(k-1)$  serbestlik dereceli  $\chi^2$  dağılımı göstermekte ve  $1/Q$ 'nun beklenen değeri  $k-1 > 2$  veya  $k > 3$  olacak şekilde  $1/((k-1)-2)$  olasılıkla ters ki-kare dağılımı göstermektedir. Bu yüzden  $I^2$  istatistiğinin beklenen değeri  $\tau^2=0$  ve  $k > 3$  için,

$$E(I^2) = \frac{-2}{k-3} \cdot 100 \quad (2.47)$$

şeklinde hesaplanmaktadır. Bu olasılık, meta analizine alınan çalışma sayısına bağlı olarak en az 4 çalışma için tanımlanmaktadır. Eğer çalışma sayısı artırılırsa bu bağımlılık daha küçük olacaktır (Mittlböck, & Heinzl, 2006).  $I^2$  için %95 güven aralığının alt ve üst sınır değerleri eşitlik (2.48) ve (2.49)'de sunulmuştur.

$$A = \exp\left(0,5 \ln\left(\frac{Q}{k-1}\right) - 1,96B\right) \quad (2.48)$$

$$U = \exp\left(0,5 \ln\left(\frac{Q}{k-1}\right) + 1,96B\right) \quad (2.49)$$

Alt ve üst sınır değerlerinde yer alan  $B$  standart hata değeridir ve Cochran  $Q$  istatistiğinin çalışma sayısından büyük veya küçük olmasına bağlı olarak farklı formüllerle hesaplanmaktadır. Eğer  $Q > k$  ise eşitlik (2.50),  $Q \leq k$  ise eşitlik (2.51)'deki gibi hesaplanır.

$$B = \frac{1}{2} \left( \frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{2(k-3)}} \right) \quad (2.50)$$

$$B = \sqrt{\frac{1}{2(k-2) \left(1 - \frac{1}{3(k-2)^2}\right)}} \quad (2.51)$$

%95 güven aralığına ait alt ve üst sınır değerleri,

$$A_{I^2} = \frac{(A^2 - 1)}{A^2} \cdot 100 \quad (2.52)$$

$$U_{I^2} = \frac{(U^2 - 1)}{U^2} \cdot 100 \quad (2.53)$$

şeklinde hesaplanır.  $I^2$  istatistiğinin, alt sınır veya üst sınır değerlerinden herhangi biri 0'dan küçükse bu değere 0 değeri atanır. Eğer alt sınır 0'dan büyük ise  $I^2$  istatistiksel olarak anlamlı kabul edilir. (Borenstein ve ark., 2009).

Heterojenlik ölçüsü bir ölçek üzerinde doğal olarak 0 ile %100 arasında görülmektedir. %100'e yakın değerler söz konusu ise heterojenlik çok büyük, sıfıra yakın değerler söz konusu ise küçük heterojenlik şeklinde değerlendirilir. Higgins ve Thompson,  $I^2$  değeri için %25'in düşük, %50'nin orta ve %75'in ise yüksek heterojenliği ifade ettiği üç kategori belirlemişlerdir. Dolayısıyla, %50'nin üzerindeki değerler için heterojenliğin önemli olduğu düşünülmelidir (Higgins, Thompson, Deeks, & Altman, 2003; Mittlböck, & Heinzl, 2006).

#### 2.4. Meta Analizinde Model Seçimi

Meta analizi çalışmalarında, çalışmaların sonuçlarını birleştirmek için sabit etki ve rasgele etkiler modelleri kullanılmaktadır. Sabit etki modelinde her çalışma için ortak olduğu varsayılan sadece bir gerçek etki büyüklüğü söz konusu olduğu için "etki" terimi, rasgele etkiler modelinde ise gerçek etki büyüklüğü çalışmadan çalışmaya değişebileceği ve birden fazla gerçek etki büyüklüğü olduğu için "etkiler" terimi kullanılır.

Hangi modelin seçileceği çalışmaların ortak etki büyüklüğüne sahip olup olmadığına dair beklentimize ve analizi gerçekleştirme hedeflerimize bağlıdır. Sabit etki modeli, meta analizine alınan tüm çalışmaların gerçek etki büyüklüklerinin aynı olduğu durumlarda tercih edilmektedir. Ayrıca, çalışmanın amacı, tanımlanmış bir anakütle için ortak bir etki büyüklüğü hesaplamak ve diğer anaküteller için bir

genelleme yapmamak olduğunda da sabit etki modeli tercih edilmelidir. Bağımsız çalışan farklı araştırmacılar tarafından yapılan çalışmalar söz konusu olduğunda, çalışmaların işlevsel olarak eşdeğer olması mümkün olmadığından ortak bir etki büyüklüğü varsayılmaz ve rasgele etkiler modeli tercih edilmelidir (Borenstein ve ark., 2009).

Çalışmaların sonuçlarının birleştirilmesinde hangi modelin kullanılacağına karar verilebilmesi için çalışmalar arasındaki etkilerin heterojenliğinin de araştırılması gerekir (Takkouche, Cadarso-Suarez, & Spiegelman, 1999). Eğer çalışmalar arası varyans istatistiksel olarak farklılık gösteriyorsa sabit etki modelinin kullanımı uygun değildir (Borenstein ve ark., 2009). Rasgele etkiler modeli ve sabit etki modeli, çalışmalar arasında heterojenite olmadığı aynı sonuçları verecektir. Heterojenlik söz konusu olduğunda, sabit etki modeli yerine rastgele etkiler modeli kullanılırsa, ortak etki büyüklüğü için güven aralıkları daha geniş olacaktır (Higgins, & Green, 2008).

#### **2.4.1. Sabit Etki Modeli (Fixed Effect Model)**

Sabit etki modeli, analizdeki tüm çalışmaların altında yatan tek bir gerçek etki büyüklüğü (dolayısıyla sabit etki terimi) olduğu ve gözlenen etkilerdeki tüm farklılıkların örnekleme hatasına bağlı olduğu varsayımına dayanmaktadır. Buna sabit etki modeli ya da ortak etki modeli denir (Borenstein ve ark., 2009).

Meta analizine alınan bütün çalışmalarda aynı faktörlerin, etki büyüklüğünü etkileyeceği varsayımı nedeniyle gerçek etki büyüklüğü tüm çalışmalarda sabittir. Gerçek etki büyüklüğü  $\theta$  ile gösterilir ve tüm çalışmalardaki gerçek etki büyüklüklerinin ortalamasına eşittir (Borenstein ve ark., 2009).

Tüm çalışmalarda gerçek etki büyüklükleri aynı olduğundan, gözlenen etki büyüklüğü bir çalışmadan diğerine sadece çalışmadaki rasgele hata nedeniyle değişir. Her bir çalışmaya ait gözlenen etki büyüklüğü ( $Y_i$ ), her bir çalışmaya ait gerçek etki büyüklüğü (anakütlenin ortalaması) ile çalışmanın örnekleme hatasının toplamıdır, eşitlik (2.54)'de verilmiştir (Borenstein ve ark., 2009).

$$Y_i = \mu + \varepsilon_i \quad (2.54)$$

Meta analizinde gözlenen etki büyüklükleriyle, anakütlenin etki büyüklüğü tahmin edilmeye çalışılır. Anakütle etki büyüklüğünün gerçek değerini tahmin etmek için eşitlik (2.55)'te sunulan ağırlıklı ortalama (M) hesaplanır (Borenstein ve ark., 2009).

$$M = \frac{\sum_{i=1}^k w_i Y_i}{\sum_{i=1}^k w_i} \quad (2.55)$$

Burada ( $Y_i$ ), i. çalışmadaki gözlenen etki büyüklüğü değeridir ve  $N(\theta, v_i)$  ile normal dağılmaktadır. k, meta analizine alınan çalışma sayısı,  $w_i$  ise her bir çalışmanın ağırlığını göstermektedir. Her bir çalışmanın ağırlığı, o çalışmaya ait varyansın (çalışma içi varyans) tersi olarak ifade edilir ( $w_i = \frac{1}{v_{Y_i}}$ ) (Borenstein ve ark., 2009).

$$v_M = \frac{1}{\sum_{i=1}^k w_i} \quad (2.56)$$

$$SE_M = \sqrt{v_M} \quad (2.57)$$

%95 Güven aralığı ise,

$$\%95 \text{ G. A.} = M \pm 1,96 SE_M \quad (2.58)$$

olarak ifade edilmektedir. Gerçek etki büyüklüğünün sıfır olduğu (fark için) ya da bir olduğu (oran için) sıfır hipotezini test etmek için Z değeri,

$$Z = \frac{M}{SE_M}$$

kullanılarak hesaplanabilir.

#### 2.4.2. Rasgele Etkiler Modeli (Random Effects Model)

Rastgele etkiler modeli ortak etki büyüklüğü varsayımına dayanmaz. Rasgele etkiler modelinde etki büyüklüğünü etkileyebilecek ve hatta etki büyüklüğünde değişikliklere neden olacak faktörler olduğu hesaba katılır (Borenstein ve ark., 2009).

Rastgele etkiler modeli ile yapılan meta analizinde gerçek etkinin normal dağıldığı varsayılmaktadır. Rasgele etkiler modelinde, her bir çalışma için gözlenen etki büyüklüğü ( $Y_i$ ), gerçek etki büyüklüğü ( $\theta_i$ ) ve anakütle ortalaması ( $\mu$ ) arasındaki

değişim  $((\theta_i - \mu) = \xi_i)$  ile çalışmanın gerçek ve gözlenen etki büyüklüğü arasındaki değişimin  $((Y_i - \theta_i) = \varepsilon_i)$  toplamıdır. Eşitlik (2.59)'deki gibi verilmiştir (Borenstein ve ark., 2009).

$$Y_i = \mu + \xi_i + \varepsilon_i \quad (2.59)$$

Her bir çalışmaya ait gerçek etki büyüklüğünün  $(\theta_i)$  anakütle ortalamasından  $(\mu)$  uzaklığı, çalışmalar arasındaki gerçek etki büyüklüklerinin dağılımının standart sapmasından kaynaklanmaktadır. Standart sapma değeri, çalışmalar arası değişkenlikle ilgili hataya ait standart sapma değeridir ve  $\tau$  ile gösterilir. Her bir çalışma için gerçek etki büyüklüğü biliniyor ve varyansı hesaplanabiliyorsa, hesaplanan bu varyans çalışmalar arası varyansı  $(\tau^2)$  göstermektedir (Borenstein ve ark., 2009).

Her bir çalışmaya ait gerçek etki büyüklüğünün  $(\theta_i)$  her bir çalışmaya ait gözlenen etki büyüklüğünden  $(Y_i)$  uzaklığı, her bir çalışmadan gözlenen etki büyüklüğünün varyansına  $(v_{Y_i})$  bağlıdır (Borenstein ve ark., 2009).

Her iki etki modelinde de her bir çalışma varyansın tersi ile ağırlıklandırılır. Sabit etki modelinde varyans, çalışma içi varyans olarak değerlendirilirken, rasgele etkiler modelinde çalışma içi ve çalışmalar arası varyansın toplamı olarak ifade edilmektedir. Bu nedenle sabit etki modeli rasgele etkiler modeline göre daha dar güven aralığına sahiptir. Rasgele etkiler modelinde, her bir çalışmanın ağırlığı sabit etki modelindeki ile aynı olmakla birlikte sadece gösterimde “\*” ile vurgulanmaktadır (Borenstein ve ark., 2009).

$$w_i^* = \frac{1}{v_{Y_i}^*} \quad (2.60)$$

Burada  $(v_{Y_i}^*)$ , i. çalışmanın çalışma içi varyans ile çalışmalar arası varyans tahmininin toplamıyla elde edilir.

$$v_{Y_i}^* = v_{Y_i} + \tau^2 \quad (2.61)$$



Ağırlıklı ortalama ( $M^*$ ) ise eşitlik (2.62) ile hesaplanır.

$$M^* = \frac{\sum_{i=1}^k w_i^* Y_i}{\sum_{i=1}^k w_i^*} \quad (2.62)$$

$$v_{M^*} = \frac{1}{\sum_{i=1}^k w_i^*} \quad (2.63)$$

$$SE_{M^*} = \sqrt{v_{M^*}} \quad (2.64)$$

ve %95 güven aralığı ise,

$$\%95 \text{ G.A.} = M^* \pm 1,96 SE_{M^*} \quad (2.65)$$

olarak ifade edilmektedir (Borenstein ve ark., 2009). Ortalama etkinin sıfır olduğu sıfır hipotezini test etmek için Z değeri,

$$Z^* = \frac{M^*}{SE_{M^*}} \quad (2.66)$$

kullanılarak hesaplanabilir.

## 2.5. İkili Verilerin Özet İstatistiklerinin Birleştirilmesinde Kullanılan Yöntemler

İkili sonuçların birleştirilmesinde yaygın olarak kullanılan üç sabit etki modeli için (Mantel-Haenszel, Peto ve varyansa dayalı yöntem) ve bir rassal etkiler modeli için (DerSimonian ve Laird) olmak üzere dört meta analizi yöntemi vardır. Peto yöntemi yalnızca odds oranlarını bir araya getirirken, diğer üç yöntem odds oranlarını, risk oranlarını ve risk farklılıklarını bir araya getirebilir. 2x2 düzeninde verilen tablolarda bulunan gözlerden biri sıfır değerini içeriyorsa, 2x2 tablonun her hücresine 0,5 eklenir (Higgins, & Green, 2008).

### 2.5.1. Mantel Haenszel Yöntemi

Mantel Haenszel (Mantel, & Haenszel, 1959) yöntemi sabit etki varsayımı altında 2x2 düzeninde verilmiş veri kümesi şeklindeki çalışma sonuçlarını birleştirmek için yaygın olarak kullanılan bir yöntemdir. Veri az olduğunda varyans yönteminde kullanılan etki büyüklüklerinin standart hatalarının değerleri zayıf olabilir. Mantel Haenszel yöntemleri, alternatif ağırlıklı bir şema kullanır. Veri az olduğunda daha

güçlü olduğu görülür ve bu nedenle varyans yöntemine tercih edilebilir. Diğer durumlarda varyans yöntemine benzer tahminler verir (Higgins, & Green, 2008).

Bu yöntemde özet istatistiği olarak odds oranı tercih edilmektedir (Petitti, 2000). Mantel Haenszel yöntemi için Tablo 1 dikkate alındığında birleştirilmiş odds oranı

$$OR_{MH} = \frac{\sum_{i=1}^k OR_i \times W_i}{\sum_{i=1}^k W_i} \quad (2.67)$$

$$Var_i = \frac{n_i}{b_i \times c_i} \quad (2.68)$$

$$W_i = \frac{1}{Var_i} \quad (2.69)$$

şeklinde elde edilir. Burada,  $OR_i$  her bir çalışmanın özet istatistiği,  $k$  çalışma sayısı ve  $W_i$  her bir çalışmanın ağırlığını göstermektedir (Borenstein ve ark., 2009). Odds oranına göre %95 güven aralığı,

$$VarOR_{MH} = \left( \frac{\sum_{i=1}^k F}{2 \times (\sum_{i=1}^k R)^2} \right) + \left[ \frac{\sum_{i=1}^k G}{(2 \times \sum_{i=1}^k R \times \sum_{i=1}^k S)} \right] + \left( \frac{\sum_{i=1}^k H}{2 \times (\sum_{i=1}^k S)^2} \right) \quad (2.70)$$

$$F = a_i d_i \frac{(a_i + d_i)}{n_i^2} \quad (2.71)$$

$$G = \frac{[a_i d_i (b_i + c_i)] + [b_i c_i (a_i + d_i)]}{n_i^2} \quad (2.72)$$

$$H = \frac{b_i c_i (b_i + c_i)}{n_i^2} \quad (2.73)$$

$$R = a_i d_i / n_i \quad (2.74)$$

$$S = b_i c_i / n_i \quad (2.75)$$

olmak üzere eşitlik (2.76)'daki gibi ifade edilir (Petitti, 2000).

$$\%95 G.A = e^{\ln OR_{MH} \pm 1,96 \sqrt{VarOR_{MH}}} \quad (2.76)$$

### 2.5.2. Peto Yöntemi

Peto yöntemi, Mantel Haenszel yönteminin bir modifikasyonudur. Mantel Haenszel yöntemi gibi çoğunlukla sabit etki varsayımı altında kullanılır. Özet istatistiği odds oranı olduğu zaman verileri birleştirmek için kullanılan alternatif bir yöntem olup Mantel Haenszel yöntemine benzer ancak hesaplaması daha kolaydır (Petitti, 2000). Peto yöntemi ile odds oranı ve %95 güven aralığını hesaplamak için Tablo 1 dikkate alındığında aşağıdaki yol izlenir:

- I. Her bir çalışmanın hasta grubundaki olayların beklenen sayısı (2.77) eşitliğindeki gibidir.

$$E_i = \frac{(e_i \times g_i)}{n_i} \quad (2.77)$$

- II. Her bir çalışmanın hasta grubundaki olayların gözlenen sayısı ile beklenen sayısı arasındaki fark  $Fark_i = O_i - E_i$ 'dir.
- III. Her bir çalışmanın gözlenen ile beklenen olay sayısı farkının varyansı eşitlik (2.78) ile tahmin edilir.

$$Var_i = \frac{(E_i \times f_i \times h_i)}{n_i(n_i - 1)} \quad (2.78)$$

- IV. Gözlenen ve beklenen fark değerlerinin toplamı  $Toplam = \sum_{i=1}^k (O_i - E_i)$ 'dir.
- V. Varyans toplamları eşitlik (2.79)'deki gibidir.

$$Var_{Top} = \sum_{i=1}^k Var_i \quad (2.79)$$

- VI. Toplam fark değerinin toplam varyansa bölünmesiyle birleştirilmiş odds oranının doğal logaritması, eşitlik (2.80) ile tahmin edilir.

$$\ln OR_{Peto} = \frac{\sum_{i=1}^k (O_i - E_i)}{\sum_{i=1}^k Var_i} \quad (2.80)$$

- VII.  $\ln OR_{Peto}$ 'nin üstel değeri alınarak birleştirilmiş odds oranı  $OR_{Peto} = e^{\ln OR_{Peto}}$  ile tahmin edilir.
- VIII. Odds oranının %95 güven aralığı tahmini eşitlik (2.81)'deki gibidir (Petitti, 2000).

$$\%95 G.A = e^{\ln OR_{Peto} \pm 1,96 \sqrt{\sum_{i=1}^k Var_i}} \quad (2.81)$$

### 2.5.3. Genel Varyansa Dayalı Yöntem

Mantel-Haenszel ve Peto yöntemleri genellikle odds oranlarının tahmin edilmesinde kullanılmaktadır. Risk farkı ve relatif risk özet istatistiklerinin birleştirilmesinde varyansa dayalı yöntemin kullanılmasının daha uygun olacağı ifade edilmiştir (Petitti, 2000).

Risk farkının ve %95 güven aralığının tahmin edilmesinde aşağıdaki yol izlenir:

- I. Her bir çalışmanın risk farkı  $RD_i = P_{H_i} - P_{K_i}$ 'dir.
- II. Her çalışmaya verilecek olan ağırlıklar hesaplanarak, risk farkı eşitlik (2.84) ile tahmin edilir.

$$W_i = \frac{1}{Var_i} \quad (2.82)$$

$$Var_i = \frac{(g_i \times h_i)}{(e_i \times f_i \times n_i)} \quad (2.83)$$

$$RD_{GV} = \frac{\sum_{i=1}^k (W_i \times RD_i)}{\sum_{i=1}^k W_i} \quad (2.84)$$

- III. Risk farkı kestiriminin %95 güven aralığı,

$$Var_{GV} = \frac{1}{\sum_{i=1}^k W_i} \quad (2.85)$$

olmak üzere,

$$\%95 \text{ G.A} = RD_{GV} \pm 1,96\sqrt{Var_{GV}} \quad (2.86)$$

şeklindedir.

Relatif risk için güven aralığının tahmin edilmesinde aşağıdaki yol izlenir:

- I. Hasta ve kontrol grubuna ilişkin olayların oranı logaritmik dönüşümünden sonra genel varyansa dayalı yöntem uygulanabilir. İlk önce her bir çalışmanın logaritmik relatif riskleri eşitlik (2.87)'daki gibi hesaplanır.

$$\ln RR_i = \ln \left( \frac{P_{H_i}}{P_{K_i}} \right) \quad (2.87)$$

- II. Logaritmik relatif risk ve varyansı eşitlik (2.88) ile tahmin edilir.

$$Var(\ln RR_i) = \frac{(1 - P_{H_i})}{(n_{H_i} \times P_{H_i})} + \frac{(1 - P_{K_i})}{(n_{K_i} \times P_{K_i})} \quad (2.88)$$

Burada,

$$\ln RR_{GV} = \frac{\sum_{i=1}^k (W_i \times \ln RR_i)}{\sum_{i=1}^k W_i} \quad (2.89)$$

ve

$$W_i = \frac{1}{Var(\ln RR_i)} \quad (2.90)$$

şeklindedir.

III. Logaritmik relatif riskin %95 güven aralığı

$$Var_{GV} = \frac{1}{\sum_{i=1}^k W_i} \quad (2.91)$$

olmak üzere, eşitlik (2.92)'de verildiği gibidir (Petitti, 2000).

$$\%95 G.A = e^{\ln RR_{GV} \pm 1,96\sqrt{Var_{GV}}} \quad (2.92)$$

#### 2.5.4. DerSimonian-Laird Yöntemi

DerSimonian Laird yöntemi (1986) rastgele etkiler modeline dayanmaktadır. Odds oranı eşitlik (2.93) ile hesaplanır.

$$\ln OR_{D\&L} = \frac{\sum_{i=1}^k (W_i^* \times \ln OR_i)}{\sum_{i=1}^k W_i^*} \quad (2.93)$$

Burada,

$OR_{D\&L}$ , DerSimonian-Laird birleştirilmiş odds oranı,

$W_i^*$ , i. çalışmanın DerSimonian-Laird ağırlık faktörü ve  $OR_i$  ise, i. çalışmanın odds oranıdır. Çalışma içi ve çalışmalar arası varyansı içeren  $W_i^*$  ağırlık faktörü,

$$W_i = \frac{1}{Var_i} \quad (2.94)$$

olmak üzere,

$$W_i^* = \frac{1}{\left[ D + \left( 1/W_i \right) \right]} \quad (2.95)$$

ile tahmin edilir. Buradaki  $i$ . çalışmanın varyansı ( $Var_i$ ), Mantel-Haenszel yöntemi kullanılarak tahmin edilir. Rastgele etki modeli hem çalışma içi hem de çalışmalar arası varyansı içerdiği için, çalışmalar arası varyans olan  $D$  değeri eşitlik (2.96)'deki gibi hesaplanır.

$$D = maks \left\{ 0, \frac{[Q - (S - 1)] \times \sum_{i=1}^k W_i}{\left[ \left( \sum_{i=1}^k W_i \right)^2 - \sum_{i=1}^k W_i^2 \right]} \right\} \quad (2.96)$$

Buradaki  $S$ , toplam çalışma olup çalışma içi sonuçların homojenlik testi olan ve  $(k-1)$  serbestlik dereceli  $\chi^2$  dağılımına sahip  $Q$  istatistiği,

$$Q = \sum_{i=1}^k W_i (\ln OR_i - \ln OR_{D\&L})^2 \quad (2.97)$$

eşitliği ile hesaplanır.  $D$  değerinin sıfırdan büyük bir değer çıkması, çalışmaların homojen olmadığını göstermektedir. Sabit etki modelinde bu etkinin de ağırlığa dâhil edilmesi gerekmektedir. Eğer  $D$ , sıfır ya da negatif değer alıyorsa birleştirilmiş odds oranı sabit etki modeliyle aynı sonucu verecektir. %95 güven aralığı,

$$Var_{GV}^* = \sum_{i=1}^k W_i^* \quad (2.98)$$

olmak üzere,

$$\%95 G.A = e^{\ln OR_{D\&L} \pm 1,96 \sqrt{Var_{GV}^*}} \quad (2.99)$$

şeklinde ve daha geniş olacaktır (Petitti, 2000).

## 2.6. Meta Analizde Yanlılık

Meta analizi, ilgili literatürün sistematik incelemelerine dayanır (Sterne, & Harbord, 2004), dahil edilen çalışmaların matematiksel olarak bir sentezini sağlar. Eğer meta analizine dahil edilen çalışmalar ilgili tüm çalışmaların yanlı bir örnekleymişse, meta analizi tarafından hesaplanan ortalama etki de bu yanlılığı yansıtır (Borenstein ve ark., 2009). Dolayısıyla ne kadar sistematik ve kapsamlı olursa olsun, meta analizi sonuçlarının geçerliliği tehdit altındadır (Rothstein ve ark., 2005). Meta analizinin geçerliliği büyük ölçüde analize dahil edilen çalışmaların verilerinin

sonuçlarına bağlıdır ve bu nedenle olabildiğince tarafsız ve eksiksiz bir şekilde ilgili çalışmanın tanımlanmasını gerektirir.

Meta analizinde meydana gelen yanlılık bazı durumlarda özellikle sağlıklı ilgili araştırmalarda, yeni bir tedavinin yararının fazla, etkisiz veya tehlikeli bir tedavinin yanlış bir şekilde güvenli ve etkili olarak tahmin edilmesine yol açacaktır. Bu durum, hastanın zarar görmesi gibi ciddi sonuçlar doğurabilir. Bir çalışmanın yayınlanmaması ile çalışmanın başlangıcından sonuna giden yolda kullanılan tüm değerli kaynaklar boşa harcanmış olur ve bunun yanında özellikle insan gönüllülerin dahil olduğu çalışma sonuçlarının yayınlanmaması etik değildir (Rothstein ve ark., 2005).

Yanlılık sorunu, editörler araştırmacıların gönderdiği sonuçların istatistiksel önemine bakılmaksızın tüm iyi yapılmış çalışmalarını kabul ettiklerinde, dergi politikalarında olumsuz sonuçların yayınlanmasında gecikmeye neden olan durumlar değiştirildiğinde büyük ölçüde önlenilebilecektir (Whitehead, 2002). Her bir çalışmaya başlamadan önce ya da en azından sonuçlarını elde etmeden önce kayıt edilmesi de meta analizinde yanlılığı önlemek için önerilen bir stratejidir. Böylece bu kayıt, sonraki meta analizleri için tarafsız bir örnekleme çerçevesi oluşturacaktır (Rothstein ve ark., 2005). Çalışmaların başlangıçta kaydedilmesiyle, araştırmacılar en azından ilgili çalışmaların ne zaman yapıldığını bilecek ve çalışma sonuçlarını elde etmek için sorumlu araştırmacıları takip edebilecekler ve iletişim kurabileceklerdir. Kayıtlar, diğer araştırmacıları devam eden çalışmalar hakkında bilgilendirerek gereksiz tekrarların ve bu tekrarlardan kaynaklanan yanlılığın önlenmesine de yardımcı olur. İdeal olarak tercih edilen eksiksiz bir liste olsa da tüm çalışmaların kaydedilmesi sağlanamasa dahi kayıtlar, meta analizleri için tarafsız bir örnekleme çerçevesi sağlayabilir. Yani, meta analizlerinde sonuçların yönüne ve büyüklüğüne göre çalışma sonuçlarının seçilerek kullanılması nedeniyle ortaya çıkabilecek yanlılığı önleme açısından bakıldığında tam bir kayıt listesi istenen bir durumdur fakat şart değildir. Kayıtlar, bir çalışmanın sonuçlarını bilmeden önce gerçekleştiği için yanlılıktan kaçınılmış olunur.

Sutton ve arkadaşları (2000) meta analizinde güvenilir testler tarafından yanlılık tespit edildiğinde ya da yanlılıktan şüphelenildiğinde bir araştırmacı için

birkaç seçenek olduğunu belirtmiştir. Seçeneklerden biri, meta analizine dahil edilmemiş çalışmaları aramaya geri dönmektir. Diğer bir seçenek, araştırmacının meta analizine devam etmesi ve sonuçları yorumlarken dikkatli olması, ortaya çıkarılan yanlılığı hesaba katmasıdır. Son seçenek ise, kullanılmalarının tartışmalı olduğu, yanlılığın etkileri için düzeltilmiş sonuçlar sağlayan istatistiksel analizlerdir. Sutton ve arkadaşları (2000), meta analizinde düzeltme yöntemlerinin yalnızca bir duyarlılık analizi olarak kullanılması gerektiğini, araştırmacının sonuçların farklı olup olmadığını görmek için düzeltmeli ve düzeltmesiz meta analizi yaparak bir duyarlılık analizi yapabileceğini belirtmiştir. Jin ve arkadaşları (2014) meta analizinde güvenilir testler tarafından yanlılık tespit edildiğinde ya da yanlılıktan şüphelenildiğinde duyarlılık analizi yapmak için Copas seçim modelinin kullanılmasını önermiştir. Benzer şekilde Rothstein ve arkadaşları da (2005) araştırmacıların her zaman meta analizinde yanlılığın varlığını kontrol etmesini ve eğer yanlılık söz konusu ise eksik çalışmaların potansiyel etkisini değerlendirmek için bir duyarlılık analizi yapması gerektiğini belirtmiştir. Wang ve Liu (2016)'nin çalışmasında ise istatistiksel test ile yanlılık gözlemlenmesi durumunda örneklem hacmi küçük olan çalışmalar meta analizinde yanlılığa neden olabileceğinden, örneklem hacmi küçük olan bazı çalışmaların, meta analizinden çıkarılması önerilmektedir.

### **2.6.1. Meta Analizde Yanlılığa Neden Olabilen Yanlılık Çeşitleri**

Etki büyüklüğünde bir sapmaya yol açabilen yanlılıklar şunlardır: yayın yanlılığı, dil yanlılığı, ulaşılabilirlik yanlılığı, maliyet yanlılığı, benzerlik yanlılığı, sonuç yanlılığı, tekrarlama yanlılığı, atıf yanlılığı, gri literatür yanlılığı, gecikme yanlılığı, kesme yanlılığı, ülke yanlılığı, medya dikkat yanlılığı, veri tabanı yanlılığı, yayın yeri yanlılığı, bütün yayın yanlılığı (Borenstein ve ark., 2009; Easterbrook, Berlin, Gopalan, & Matthews, 1991; Egger, Smith, Schneider, & Minder, 1997; Ioannidis, 1998; Rothstein ve ark., 2005; Song ve ark., 2010; Stern, & Simes, 1997; Toews ve ark., 2017). Tüm bu yanlılıkların bulunduğu literatür, tamamlanmış çalışmaların anakütlesini temsil etmeyecektir; dolayısıyla bu yanlılıkların tümü, bir meta analizin geçerliliğini aynı şekilde tehdit eder (Rothstein ve ark., 2005). Literatürde, bahsi geçen yanlılıkların tümüne atıfta bulunmak için dar ve basit anlamlarının yanında “yayımla yanlılığı” (Song ve ark., 2010), “yayın yanlılığı”



(Rothstein ve ark., 2005), “ulaşılabilirlik yanlılığı”, “seçim yanlılığı” terimleri de kullanılmaktadır (Hunter, & Schmidt, 2004). “Yayın ve ilgili yanlılıklar” olarak da yaygın şekilde kullanıldığı görülmüştür. Bu çalışmada, bahsi geçen bütün yanlılıkları kapsaması adına “yanlılık” teriminin kullanılması tercih edilmiştir.

Meta analizine dahil edilecek çalışmaların arama sürecini etkileyen yanlılıklar:

#### ***Yayın yanlılığı (Publication bias)***

Yayın yanlılığı, yayınlanmış literatürde ortaya çıkan araştırmanın, tamamlanmış çalışmaların anakütlesinin temsilcisi olmadığını ortaya koyan terimdir. Yayın yanlılığı, sonuçların yönü ve istatistiksel önemine bağlı olarak çalışmaların yayınlanması veya yayınlanmaması olarak tanımlanmıştır (Rothstein ve ark., 2005). Tamamlanmış çalışmalar olumsuz sonuçlar, ilgi eksikliği, amaçlanan veya devam eden bir makale ve örneklem hacmi problemleri gibi nedenlerle yayınlanamayabilirler. Nispeten yüksek etki büyüklükleri bildiren, sıfır hipotezini reddeden sonuçlarla yapılan çalışmaların yayınlanma olasılığı daha yüksektir. Yayınlanan çalışmaların bir meta analizine girme olasılığı daha yüksek olduğundan, literatürdeki herhangi bir yanlılığın meta analizine de yansması muhtemeldir (Borenstein ve ark., 2009).

#### ***Yayılma yanlılığı (Dissemination bias)***

Bir çalışmanın sonuçlarının yayılma profili, bulgularının yönüne veya gücüne bağlı olduğunda oluşur. Yayılma profili, araştırma sonuçlarının erişilebilirliği veya araştırma bulgularının potansiyel kullanıcılar tarafından tanımlanması olasılığı olarak tanımlanmaktadır. Yayılma profilinin spektrumu, araştırmanın yayınlanıp yayınlanmadığına, ne zaman, nerede ve nasıl yayınlandığına göre, tamamen erişilemez durumdan kolay erişilebilirliğe kadar uzanmaktadır (Song ve ark., 2010).

#### ***Dil yanlılığı (Language bias)***

Dil yanlılığı, yayın dillerinin çalışma sonuçlarının yönüne ve gücüne bağlı olduğu zaman oluşur (Borenstein ve ark., 2009; Song ve ark., 2010). İngilizce yayınlanan çalışmaların seçilerek araştırmaya dahil edilmesi ya da araştırmacıların bildiği dillerde literatürü taraması gibi nedenlerle ortaya çıkar. Araştırmacıların prestijli İngilizce dergilerde önemli sonuçlar yayınlamak için daha istekli olmalarına rağmen, anlamlı olmayan çalışmaları kendi anadillerinde yayınlama olasılıkları daha

yüksektir. İngilizce veri tabanları ve dergilerin, aynı zamanda İngilizce dilinin yaygın kullanılması, ya da bu veri tabanları ve dergilerin araştırılmasının daha mümkün olması nedeni ile de istatistiksel olarak anlamlı çalışmaların aşırı örneklenmesi meydana gelebilmektedir (Rothstein ve ark., 2005). Özetle, İngilizce dışındaki dillerde yayınlanan çalışmalar, İngilizce yayınlanmış olanlardan farklı sonuçlar verebilmektedir. Bu nedenle, bir araştırmanın dil yanlılığı içermemesini sağlamanın en iyi yolu, meta analizine dahil edilecek tüm çalışmaları yayınladıkları dilden bağımsız olarak aramak ve dahil etmektir.

### ***Sonuç yanlılığı (Outcome bias)***

Yazarın, birincil sonuçları seçerek raporlaması, ancak sonuçların yönüne ve istatistiksel önemine bağlı olarak diğer sonuçları raporlamaması anlamına gelen yanlılıktır. Bazen yayına sunulan makalede mecburi olarak tüm analizlerin sonuçlarını sunacak yer ve imkân olmayabilir. Bu gibi durumlarda sonuçların seçilmesi kaçınılmazdır ve yalnızca en olumlu, heyecan verici veya önemli sonuçların sunulması ile yanlılığa sebebiyet verilir (Rothstein ve ark., 2005; Song ve ark., 2010).

### ***Tekrarlama yanlılığı (Duplication bias)***

Tekrarlayan yayın, benzer makalelerin birden fazla dergiye sunulması veya aynı verilerin iki veya daha fazla dergide yayımlanması olarak tanımlanır (Song ve ark., 2010). Tekrarlama yanlılığı ise birden çok yayının sonucu olarak çalışmaların birden fazla kez meta analizine dahil edilmesiyle ortaya çıkan yanlılıktır (Rothstein ve ark., 2005).

Uzun vadeli çalışmaların sonuçlarını bildirmek veya belirli bir kitle için daha erişilebilir olan formatta sonuçlar sunmak gibi nedenlerle bir çalışma birden fazla yayımlanabilir. Yayınlar, bir çalışmanın benzer bir bölümünü, hipotezlerini, yöntemlerini, sonuçlarını ve / veya tartışmasını paylaşabilir (Song ve ark., 2010). Pratikte ise yazarlar değiştiği, farklı sayıda hasta bildirildiği veya deneme tasarımı farklı olarak tanımlandığı için aynı araştırmanın birden fazla yayını ayırt etmek zordur ve yanlılık meydana gelir.

İstatistiksel olarak anlamlı sonuçlara sahip çalışmaların, anlamlı olmayanlara göre birden fazla yayımlanması, yüksek atıf etki faktörü olan dergilerde yayımlanması

olasılığı daha yüksektir (Easterbrook ve ark., 1991). Bu da bir etki büyüklüğünün fazla tahmin edilmesine yol açabilir.

Tekrarlanan yayınlar "açık" veya "gizli" olarak sınıflandırılabilir. Açık tekrarlanan yayın, orijinal raporların uygun çapraz referansı ile bir çalışmadan verilerin yeniden analiz edilmesi olarak tanımlanır. Gizli tekrarlanan yayın, aynı verilerin farklı yerlerde veya farklı zamanlarda önceki veya paralel bir yayına yeterli atıfta bulunmadan yayınlanmasıdır. Yanlılık, gizli tekrarlanan yayın nedeniyle aynı çalışmadan olan verilerin birden fazla kez eklenmesiyle meta analizine sokulabilir (Song ve ark., 2010).

Meta analizine dahil edilecek uygun olan her çalışmanın yalnızca bir kez dahil edilmesini sağlamak önemlidir. Aynı verilerin tekrarlanan yayınlanmasına ilişkin yanlılıktan kaçınmak için, herhangi bir yayın dahil edilen verilerin kaynaklarını açıkça belirtmeli ve aynı verilerin önceki yayınlarına atıfta bulunmalıdır. Birden fazla yayınlanan çalışma türleri arasında açıkça belirtilen bir çapraz referans olmalıdır (Rothstein ve ark., 2005).

#### ***Atıf yanlılığı (Citation bias)***

Atıf yanlılığı, sonuçların niteliğine ve yönüne bağlı olarak araştırma bulgularının alıntılanması veya alıntılanmaması nedeni ile meydana gelen yanlılıktır. İlgili konu hakkında bilgi içeren makalelerin referans listelerinin incelenmesi, ilgili olabilecek ilave makaleleri belirlemek için kullanılabilir. Bu yaklaşımdaki sorun, önceki çalışmalara atıf yapmanın olası motivasyonlarından olan ilgilileri ikna etme ve yazarın kendi bakış açısını haklı çıkarma arzusu ile atıfta bulunulması sonucu yanlı bir çalışma örneği üretilmesidir (Rothstein ve ark., 2005).

#### ***Gri literatür yanlılığı (Gray literature bias)***

Üçüncü Uluslararası Gri literatür Konferansı (Third International Conference on Grey Literature), gri literatürü “devlet, akademi, iş dünyası ve sanayinin her düzeyinde basılı ve elektronik formatlarda üretilen, ancak ticari yayıncılar tarafından kontrol edilmeyen” olarak tanımlamıştır (Rothstein ve ark., 2005; Song ve ark., 2010). Gri literatüre örnek olarak konferans özetleri, araştırma raporları, kitap bölümleri, yayınlanmamış veriler, web sayfaları, veri arşivleri, tezler, politika belgeleri, teknik

raporlar, ödevler, toplantı bildirimleri, devam eden araştırma veri tabanları ve kişisel yazışmalar verilebilir.

Gri literatür yanlılığı, daha az dikkate değer çalışma bulgularının hakemli dergiler dışındaki yayınlarda daha fazla yayınlanması anlamına gelen yanlılıktır. Dergi makalelerinde bildirilen sonuçlar, gri literatürde sunulanlardan sistematik olarak farklı olduğunda oluşur. Yayınlanmış literatür, karşılık gelen gri literatüre göre daha olumlu olma, daha büyük bir etki büyüklüğü bildirme eğilimindedir (Song ve ark., 2010). Gri literatürün meta analizlerine dahil edilmesi etki büyüklüklerini azaltır. Gri literatürün bir meta analizine dahil edilmesi, yanlılığı en aza indirmenin önemli bir yoludur.

#### ***Ulaşılabilirlik yanlılığı (Availability bias)***

Araştırmaya kolaylıkla erişilebilen çalışmaların seçilerek dahil edilmesi sonucu ortaya çıkabilen yanlılıktır (Rothstein ve ark., 2005). Örnek olarak, eğer veri sadece belirli deneme türlerinden elde edilirse ulaşılabilirlik yanlılığı meydana gelir.

#### ***Maliyet yanlılığı (Cost bias)***

Araştırmaya ücretsiz veya düşük maliyetli çalışmaların seçilerek dahil edilmesi ile ortaya çıkabilen yanlılıktır (Rothstein ve ark., 2005). Düşük maliyet ile ulaşılabilen ve araştırmaya dahil edilen çalışmalar ilgili konu ile alakalı literatürün temsili olmayabilir.

#### ***Benzerlik yanlılığı (Familiarity bias)***

Yalnızca kendi disiplininde seçilen çalışmaların dahil edilmesi ile ortaya çıkabilen yanlılıktır (Rothstein ve ark., 2005).

#### ***Gecikme yanlılığı (Time-lag bias)***

Yanlılığın bir biçimi de gecikme yanlılığıdır. Gecikme yanlılığı, bir çalışmayı tamamlamak ve yayınlamak için geçen sürenin çalışma sonuçlarından etkilendiği anlamına gelir (Ioannidis, 1998). Yani yayınlama hızı, çalışmanın sonuçlarının yönüne ve gücüne bağlıdır (Song ve ark., 2010).

Yazarlar, ilgi çekici veya popüler bir görüşü destekleyen bulguların yayınlanması konusunda daha sabırsız olabilir; editörler daha çok haber değeri taşıyan

bulguların yayınlanmasına öncelik verebilir (Toews ve ark., 2017). Bu gibi nedenle, istatistiksel olarak anlamlı (pozitif-olumlu) sonuçları olan çalışmalar, anlamsız (negatif-olumsuz) sonucu olan çalışmalardan daha erken yayınlanma eğilimindedir (Ioannidis, 1998; Stern, & Simes, 1997). Yayın zamanı ile etki büyüklüğü arasındaki ilişki incelenerek gecikme yanlılığı araştırıldığında yayınlanma tarihi ile etki büyüklüğü arasında negatif bir ilişki olduğu ortaya çıkmıştır (Jennions, & Moller, 2002).

Daha önce önemli sonuçların yayınlanması ile ortaya çıkan gecikme yanlılığının bir sonucu, zaman içinde yapılan çalışmalar tarafından bildirilen azalan etki büyüklüğü olabilir. Bu nedenle, meta analizinde bildirilen etki büyüklüğünün zamansal eğilimleri, gecikme yanlılığının varlığını gösterebilir (Song ve ark., 2010).

#### ***Kesme yanlılığı (Truncation bias)***

Bilimsel dergilerin kelime sınırlaması, araştırmaların bulgularını tam olarak raporlaması için kısıtlayıcı olabilmektedir. Raporlar, kitaplar ve tezler gibi yayınlarda sınırlamaların daha az olması ve daha çeşitli biçimlere izin verilmesinden dolayı, çalışmaların daha rasgele ve tüm bulguları rapor etmesi daha muhtemeldir. Bilimsel dergilerde olduğu gibi kelime sınırlarıyla yayınlanan çalışmalar, bulguların eksik raporlanmasına neden olabilir (Toews ve ark., 2017).

#### ***Bütün yayın yanlılığı (Full publication bias)***

Özet çalışmalar, bir meta analizine dahil edilmek üzere yapılan çalışmaları belirlemek için önemli bir kaynaktır. Bu özetler genellikle elektronik veri tabanlarına dahil edilmez. Dergilerin çalışmayı kabul etmeyeceği, sonuçların yeterince önemli olmadığı düşüncesi gibi nedenlerle araştırmacılar tarafından bilimsel toplantılara sunulan bu özetler daha sonra tam olarak yayınlanmayabilir. Bütün yayın yanlılığı, başlangıçta konferans bildirileri, bilimsel toplantılar gibi formatlarda özet biçiminde sunulan çalışmaların tam olarak yayınlanması bulgularının yönüne ve / veya gücüne bağlı olduğunda ortaya çıkar (Song ve ark., 2010).

#### ***Yayın yeri yanlılığı (Place of publication bias)***

Yayın yeri, çalışma bulgularının yönü veya gücü ile ilişkili olduğunda ortaya çıkan yanlılıktır (Rothstein ve ark., 2005). Örneğin, olumlu sonuçlara sahip olan

çalışmaların, olumsuz sonuçlara sahip çalışmalara göre daha geniş çaplı ve etkili dergilerde yayınlanması daha olası olabilir; bir dergi yayın politikası veya okuyucuların tercihi nedeniyle belirli bir hipotezle ilgili makaleleri diğer dergilerden daha fazla yayınlamaya eğilimli olabilir (Song ve ark., 2010).

#### ***Veri tabanı yanlılığı- indeksleme yanlılığı (Database bias - Indexing bias)***

Konuyla ilgili araştırmalar genellikle elektronik bibliyografik veri tabanlarının araştırılmasıyla başlar. Fakat, MEDLINE gibi literatür veri tabanları bir konuda yayınlanan tüm çalışmaları içermeyebilir ve endeksleyemeyebilir. Literatür araştırmasının endeksli çalışmaların sonuçlarının sistematik olarak endekslenmemiş çalışmalardan farklı olduğu bir veri tabanına dayandığında ortaya çıkabilen yanlılıktır (Song ve ark., 2010).

#### ***Medya dikkat yanlılığı (Media attention bias)***

Genel nüfus, bilim ve tıp alanındaki en son gelişmeler hakkındaki bilgilerinin çoğunu popüler medyadan almaktadır. Basının bu gelişmelerin bulgularını nasıl sunduğu, halkın algısı üzerinde çok güçlü bir etkiye sahiptir. Medya dikkat yanlılığı, çarpıcı sonuçlar içeren çalışmaların gazeteler, radyo ve televizyon haberleri tarafından kapsanması daha muhtemel olduğunda ortaya çıkar (Song ve ark., 2010).

#### ***Ülke yanlılığı (Country bias)***

Farklı ülkeler arasında aynı konuda yapılan çalışmalardan elde edilen değişken sonuçların nedenlerinden biri olarak ülke yanlılığı ortaya çıkar. Bazı ülkelerden (özellikle gelişmekte olan ülkelere) olağandışı yüksek pozitif sonuç oranları bildirilebilmektedir (Song ve ark., 2010).

### **2.6.2. Yanlılıktan Kaçınmak İçin Yapılması Gerekenler**

Meta analizinde, doğru sonuçlar elde edebilmek için ilk yapılması gereken, analize dahil edilecek çalışmaların sistemli ve dikkatli bir şekilde seçilmesidir. Meta analizinde kullanılmak üzere seçilen çalışmaların oluşturduğu örneklemin, anakütleyi temsil edebilmesi için hangi çalışmaların meta analizine seçileceği, bütün kaynakların aranması önemlidir. Ne aranırsa aransın, meta analizi yazarlarının açık olması ve incelemede tam olarak hangi kaynakların arandığını ve hangi stratejileri

kullandıklarını belgelemeleri önemlidir, böylece okuyucu inceleme sonuçlarının geçerliliğini değerlendirebilir. Araştırmaların nispeten sınırlı kapsamı, incelemelerin sonuçlarının yanlış olduğu anlamına gelebilir (Rothstein ve ark., 2005).

### ***Merkezi kayıt defterlerinin aranması***

Konuyla ilgili araştırmalar genellikle elektronik bibliyografik veritabanlarının araştırılmasıyla başlar. Merkezi kayıt defterleri meta analizinde yanlışlık oluşturma riskini en aza indirmek için çalışmışlardır. Uluslararası bir kuruluş olan “Cochrane Collaboration”, “Cochrane Kontrollü Araştırmalar Merkezi Kayıt Defteri” (CENTRAL) ile en kapsamlı kayıt kaynaklarından biridir. CENTRAL, MEDLINE veya EMBASE gibi elektronik veritabanlarında endekslenmeyen randomize çalışma raporlarının yanı sıra; İngilizce dışındaki dillerde yayınlanan alıntılar, yalnızca konferans bildirilerinde yayınlanan alıntılar ve bulunması zor diğer kaynaklardan alıntılar içerir (Rothstein ve ark., 2005).

### ***Elektronik veri tabanlarında ve el ile arama***

MEDLINE, EMBASE, ERIC ve PsycINFO gibi elektronik bibliyografik veri tabanlarının araştırılması, meta analize dahil edilmesi için çalışmaların tanımlanmasında önemli bir araçtır. Bununla birlikte, bir araştırma yalnızca elektronik veri tabanlarıyla sınırlandırılmamalıdır, çünkü bu şekilde bir araştırma yanlışlığa neden olma potansiyeline sahiptir. Bütün çalışmalar elektronik veri tabanlarında yer almayabilmektedir. Elektronik araştırmaların başarısızlığının ana nedeni, bazı raporların özet olarak ve/veya elektronik veri tabanları tarafından rutin olarak endekslenmeyen dergi eklerinde yayınlanmış olmasıdır. Elektronik araştırmaların çalışmaların tüm raporlarını tespit edememesinin bir başka nedeni de bazı raporların, araştırma raporlarını indeksleme sisteminin bugünkü kadar iyi geliştirilmediği tarihlerden önce yayınlanmış olmasıdır (Rothstein ve ark., 2005).

### ***Konferans bildirilerinin aranması***

Konferans bildirileri de sistematik incelemelere dahil edilecek çalışmaları belirlemek için önemli bir kaynaktır. Bu özetler genellikle elektronik veri tabanlarına dahil edilmez ve ilgili çalışmalar yalnızca konferansın bildirileri taranarak tanımlanır. Bir konferansta bildirilen bir çalışmanın tam olarak yayınlanmaması sonuçlarının önemi ile bağlantılıdır. Bu, çalışmaların bir kısmı ve sonuçları için, mevcut tek bilgi

kaynağının bir konferans bildirisinde yayınlanan özet olabileceği anlamına gelir. Bilimsel toplantılara sunulan araştırmaların sonradan tam olarak yayınlanmamasının en yaygın nedeni, zamanın olmamasıdır. Ayrıca, araştırmacıların bir derginin çalışmalarını kabul etmeyeceklerini düşünmeleri veya yazarların sonuçların yeterince önemli olmadığını düşünmeleri de bu nedenler arasındadır (Rothstein ve ark., 2005).

### ***Araştırmacılarla iletişim kurma***

Bazı çalışmalar hiçbir zaman herhangi bir biçimde yayınlanmayacaktır ve yayınlanmamış çalışmaların, yayınlanmış çalışmalardan farklı sonuçlara sahip olabileceğine dair kanıtlar göz önüne alındığında, yanlılığı en aza indirmek için bunları tespit etmek ve inceleme sürecine dahil etmek önemlidir. Araştırmacıların, konunun uzmanları ve ilgili endüstri ile iletişime geçmesi, devam eden veya yayınlanmamış çalışmaları belirlemenin önemli bir yolu olabilir. Fakat, bir çalışma hakkında ek bilgi talebi her zaman başarılı olmamaktadır. Bu sorunun üstesinden gelmek için, bazı şirketler, araştırma kayıtları sağlayarak, devam eden ve tamamlanan çalışmaların kayıtlarını kamuya açık hale getirebilmektedir. Ancak, olası bir çalışmanın bilgisi de mutlaka bir inceleme için gerekli olan verilerin erişilebilir olacağı anlamına gelmemektedir (Rothstein ve ark., 2005).

### ***Araştırma kayıtlarının aranması***

Araştırma kayıtları, bazıları resmi olarak asla yayınlanamayabilecek, devam eden araştırmaları belirlemenin önemli bir yoludur. Yürütülen ve ilgilenilen devam eden sağlık araştırmalarını içeren birçok ulusal kayıt vardır. Ayrıca, her biri kendi ilgi alanına özgü çok sayıda yerel ve uzman kayıtları vardır. Devam eden çalışmaların belirlenmesi önemlidir, böylece bir meta analizi güncellendiğinde, bu çalışmalar olası dahil etme açısından değerlendirilebilir. Devam eden çalışmaların da tanımlanması gözden geçiren kişinin zaman gecikmesi yanlılığı olasılığını değerlendirmesine yardımcı olur (Rothstein ve ark., 2005).

### ***İnternette arama***

İnternet, devam eden ve tamamlanan araştırmalar, özellikle resmi olarak yayınlanmayan araştırmalar hakkında bir başka önemli bilgi kaynağı olabilir. Bununla birlikte, internette arama yapmak zahmetli bir iş olabilir. Genel arama motorlarının



çoğu karmaşık aramaya izin vermez ve kontrol edilecek binlerce web sitesi belirleyebilir (Rothstein ve ark., 2005).

### ***Farklı bilimler arasındaki aramadaki farklılıkları hesaba katmak***

Bir meta analizi yapılırken, arama süreci sırasında yanlılık getirme potansiyeli, farklı alanlarda daha da büyük olabilir. Tıbbi literatür ağırlıklı olarak hakemli dergilerde yayınlanır ve bunlardan bazılarına MEDLINE ve EMBASE gibi büyük, sofistike, bibliyografik veri tabanları aracılığıyla erişilebilir. Buna karşılık, sosyal bilimler literatürü çok daha çeşitlidir. Bazı araştırmalar bibliyografik veri tabanlarında indekslenir, ancak bu veri tabanlarının kapsamı ve kalitesi büyük ölçüde farklılık gösterir. Araştırma çalışmalarını tanımlamak için bu veri tabanlarının çoğunda belirli indeksleme terimlerinin eksikliği vardır (Rothstein ve ark., 2005).

### **2.6.3. Literatürde Geliştirilen Meta Analizinde Yanlılık Oluşturma Yaklaşımları**

Kale ve Nirpharake (2017), çalışmalarında yanlılığı arttırmak için, etki büyüklüğünün olumsuz tarafındaki çalışmaların %20'sini her bir veri kümesinden çıkarmıştır. Böylece başka bir 4000 verilik küme oluşturmuşlardır.

Peters ve arkadaşları (2006), çalışmalarında funnel grafiği asimetrisini iki şekilde sağlamıştır. İlk olarak p değeri büyüdükçe çalışmanın meta analizi dışında bırakılma olasılığı daha yüksek olacağından çalışma etki büyüklüğü ile ilişkili p değerine dayanarak uyarılmıştır. İkinci olarak, bir çalışma tahmininin, OR büyük olduğunda istatistiksel olarak anlamlı olma ihtimalinin daha yüksek olması nedeniyle, daha büyük OR'ler için daha az yanlılık olacağından, yanlılık çalışma etki büyüklüğüne dayanarak uyarılmıştır.

Duval ve Tweedie (2000) çalışmalarında anakütle ortalama büyüklüğü 0 olan 35 çalışmanın tek bir rasgele etki meta analizini simüle etmiştir. Yanlılık, funnel grafiğinden 5 çalışma kaldırılarak uygulanmıştır.

Sterne ve arkadaşları (2000), 5, 10, 20 ve 30 çalışma içeren "tipik" meta analizleri türetmiş, orta şiddette yanlılık, ağır yanlılık ve yanlılığın olmadığı durumlarda ağırlıklı regresyon yönteminin ve sıra korelasyon testinin yanlılığı tespit etme gücünü (duyarlılığını) değerlendirmek ve ne zaman yanlış pozitif sonuçlar

verdiklerini belirlemek için simülasyonlar yapmıştır. Yanlılık derecesi, (3.1) eşitliğinde verilen etki büyüklüğü (log OR) ve etki büyüklüğünün standart hatası (SE) arasında doğrusal bir ilişki varsayılarak tanımlanmıştır. Yazarlar yanlılık katsayılarını 0, -0,5 ve -1 (sırası ile yanlılık yok, orta şiddette yanlılık ve şiddetli yanlılık) olarak simülasyonlarını gerçekleştirmişlerdir.

$$\log(OR) = \text{Gerçek etki büyüklüğü (log OR)} + (\text{yanlılık katsayısı} \times \text{etki büyüklüğünün standart hatası (log OR)}) \quad (3.1)$$

#### **2.6.4. Meta Analizinde Yanlılık Belirlemede Kullanılan Yöntemler**

Funnel grafiği meta analizindeki yanlılığı belirlemede görsel bir değerlendirme sunarak, verilerin doğasının anlaşılmasına yardımcı olur. Funnel grafiğinde asimetri olduğunda bulguları daha objektif değerlendirme isteği ise bu grafiğin asimetrisi için istatistiksel testlerin geliştirilmesine yol açmıştır.

##### **2.6.4.1. Funnel Grafiği**

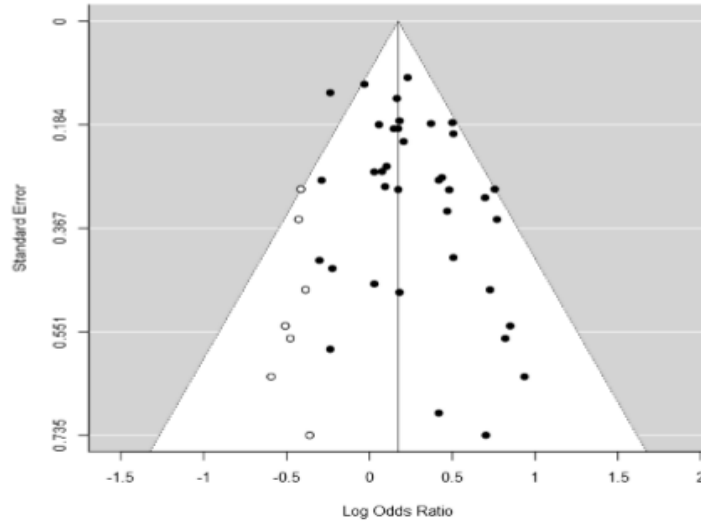
Grafik, 1984 yılında Light ve Pillemer tarafından geliştirilmiştir (Begg, & Mazumdar, 1994). Funnel grafikleri, çalışma büyüklüğünün bir ölçüsüne karşı, tek tek araştırmalardan tahmin edilen etki büyüklüğünün basit dağılım grafikleridir (Rothstein ve ark., 2005).

Bir çalışma ne kadar küçükse, sonuçların istatistiksel olarak anlamlı kabul edilmesi için gerekli etki büyüklüğü o kadar büyük olur. Ek olarak, daha büyük çalışmalara daha fazla para ve zaman yatırımı yapılması, sonuçlarının olumsuz olsa bile yüksek yöntemsel kalitede olmalarının ve yayınlanmalarının daha muhtemel olduğu anlamına gelir. Bu nedenle, bir meta analizindeki yanlılık, etki büyüklüğü ile çalışmanın örneklem hacmi arasında bir ilişkide ortaya çıkabilir (Sterne, Gavaghan, & Egger, 2000). Dolayısıyla meta analizdeki yanlılık, yatay ekseninde etki büyüklüğü, dikey ekseninde ise standart hatanın kendisi, standart hatanın tersi, örneklem hacmi, varyans, varyansın tersi ya da çalışma büyüklüğünün ağırlığının yer aldığı dağılım grafikleri olarak gösterebilir.

“Funnel Grafiği”, incelemedeki çalışmaların örneklem hacmi arttıkça, altta yatan etki büyüklüğünün tahminindeki duyarlılığın (precision) arttığı gerçeğine dayanır. Bu nedenle, dikey ekseninde çalışmanın örneklem hacmi ölçeği çizildiğinde,

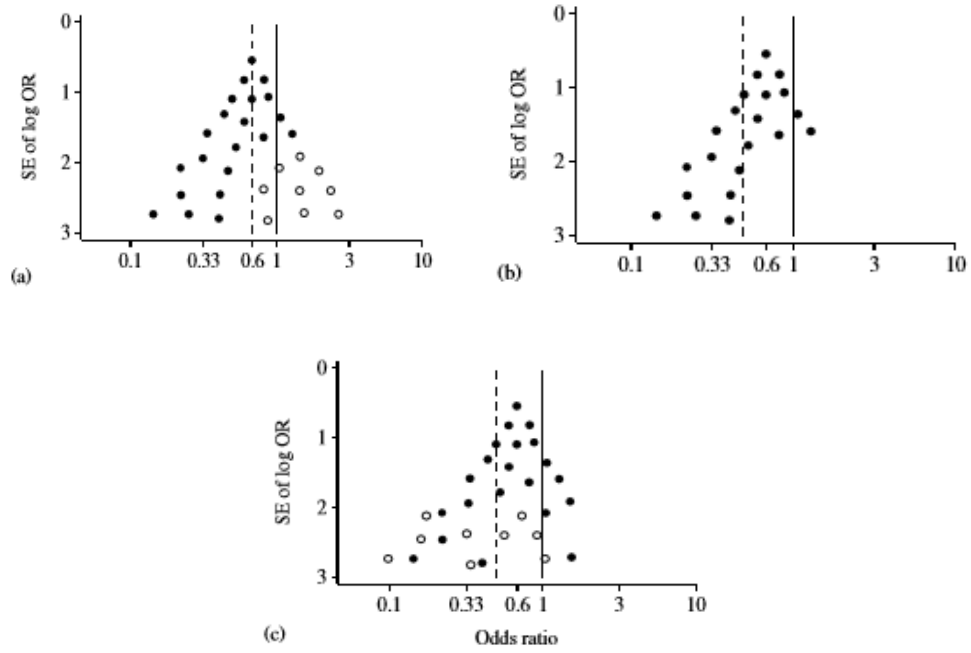
daha küçük örneklem hacmine sahip çalışmalardan elde edilen sonuçlar grafiğin dibinde yaygınlaşacak ve daha büyük örneklem hacmine sahip çalışmalar arasında dağılım daralacaktır (Egger ve ark., 1997; Sterne ve ark., 2000).

Şekil 1’de eşit büyüklükteki noktalar her bir çalışmayı ifade etmektedir. Düşey düz çizgi etki büyüklüğü değerini, eğik çizgiler ise etki büyüklüğüne ait %95 güven aralıklarını göstermektedir. Grafikte, bazı çalışmaların %95 güven aralıkları dışına düştüğü ve çalışmaların asimetrik bir şekilde dağıldığı görülmektedir. Buradaki asimetri, yanlılığın olduğu şeklinde yorumlanırken, %95 güven aralıkları dışına düşen çalışmaların ise heterojenliğe neden olduğu şeklinde yorumlar yapılabilmektedir.



Şekil 1. Funnel Grafiği (Duval, 2005)

Yanlılık yokluğunda, grafik, simetrik ters bir huniyi andırır (Şekil 2(a)). Yanlılık var ise, grafik genellikle çarpık ve asimetrik olacaktır. Örneğin, istatistiksel olarak anlamlı etkiler göstermeyen daha küçük çalışmalar yayınlanmadan kaldığı için yanlılık varsa, funnel grafiği asimetrik görünecek ve grafiğin sağ alt tarafında boşluk bırakacaktır (Şekil 2 (b)). Asimetri ne kadar belirginleşirse, yanlılık miktarının da o kadar çok olması muhtemeldir (Egger ve ark., 1997; Rothstein ve ark., 2005).



Şekil 2. Varsayımsal funnel grafikleri: (a) yanlılık yoksa funnel grafik; (b) yanlılık varlığında asimetric funnel grafik; (c) daha küçük çalışmaların düşük metodolojik kalitesinden ötürü yanlılık varlığında asimetric funnel grafik (Rothstein ve ark., 2005).

#### 2.6.4.2. Funnel Grafiği Asimetrisinin Değerlendirmesinde Kullanılan Yöntemler

Funnel grafiğinde asimetrinin olup olmadığını incelemek için çok çeşitli testler tasarlanmıştır. Funnel grafiği asimetrisinin değerlendirilmesinde kullanılan yöntemlerden Begg'in, Egger'in, Schwarzer'in, Harbord'un, Peters'in, Macaskill'in Thompson'un ve Arcsin yöntemlerinden aşağıda bahsedilmiştir. Funnel grafiği asimetrisinin değerlendirilmesinde kullanılan testler, Begg ve Mazumdar (1994) tarafından tanımlanan ve bu testten etkilenerek ortaya çıkan sıra korelasyon yöntemlerini kullanan parametrik olmayan testler ve Egger testi ve modifikasyonlarından oluşan regresyon testleri olarak iki şekilde kategorize edilebilir.

Begg ve Egger testleri, funnel grafiği asimetrisinde kullanılan klasik testler iken, diğerleri bu iki testin modifikasyonlarıdır. Yanlılık yokluğunda funnel grafiği simetrisi varsayımı, normallik varsayımı altında örneklem ortalaması, örneklem varyansından istatistiksel olarak bağımsız olduğu için, sürekli veriler söz konusu olduğunda uygundur. Bu varsayımlara dayanılarak kullanılan funnel grafiği asimetri

testlerinin kesikli veriler için kullanılması doğru değildir ve birçok yazar tarafından tartışılan bir gerçektir (Schwarzer, Carpenter, & Rücker, 2015). Özel olarak, etki büyüklüğünün log OR ile özetlendiği varsayıldığında log OR'nin varyans tahmin edicisi istatistiksel olarak log OR'nin tahminine bağlıdır. Meta analizinde yanlılık olmadığı durumda bile bu bağımlılık ikili veriler söz konusu olduğunda funnel grafiğinde bir miktar asimetriye neden olur. Niceliksel sonuçlar için tasarlanmış testler ikili sonuçlara sahip meta analizlerinde uygulandığında, potansiyel olarak funnel grafiğinde asimetri için olması gerekenden daha sık istatistiksel olarak anlamlı p değerleri rapor ederler (Schwarzer ve ark., 2015). Böylece, mevcut testlerin ikili sonuçlar için modifikasyonları ortaya çıkmıştır. Meta analizinde yanlılık belirlemede kullanılan testler için anlamlılık düzeyinin 0,10 olarak alınması önerilmiştir (Egger ve ark., 1997; Sterne ve ark., 2000).

### ***Begg'in Yöntemi***

Begg ve Mazumdar (1994), etki büyüklüğü tahminleri ile örnekleme varyansları arasındaki ilişkiyi incelemek için düzeltilmiş sıra korelasyon yöntemi önermiştir. Yanlılık varsa daha küçük çalışmalar daha büyük etkiler gösterecektir; çünkü daha küçük çalışmalardan elde edilen varyans tahminleri de büyük olacağından etki büyüklüğü ve varyans arasında pozitif bir korelasyon ortaya çıkmaktadır (Egger ve ark., 1997).

Geçerli bir sıra korelasyon testi oluşturmak için testi gerçekleştirmeden önce etki büyüklüklerini standartlaştırarak varyansları dengelemek gerekir (Begg, & Mazumdar, 1994).  $\theta_i^*$ , standartlaştırılmış etki büyüklüklerini ifade etmektedir. *i.* çalışmadan elde edilen  $\theta_i$  etki büyüklüğü tahmini (log odds oranı) ve  $v_i$  bu etki büyüklüklerinin varyansı olmak üzere;

$$\theta_i^* = (\theta_i - \bar{\theta}) / \sqrt{v_i^*} \quad (3.2)$$

(3.2)'de gösterilen eşitlikte standart etki büyüklüklerini türetmek suretiyle varyanslar dengelenir. Burada;  $\bar{\theta} = (\sum \theta_i v_i^{-1}) / \sum v_i^{-1}$  dir ve  $v_i^* = v_i - (\sum v_j^{-1})^{-1}$ ,  $(\theta_i - \bar{\theta})$ 'nın varyansıdır (Rothstein ve ark., 2005).

Test, Kendall'ın  $\theta_i^*$  ve  $v_i^*$  arasındaki sıra korelasyonunun elde edilmesine yani iki nicelikteki sıraların karşılaştırılmasına dayanır. Sıralar, sıralı değerlerdir: örneğin,  $\theta_i^*$ 'nin en büyük değeri sıralamada 1, bir sonraki en büyük değeri sıralamada 2 olur, vb. k olası çalışma varsa, o zaman  $k(k-1)/2$  olası çalışma çifti vardır. i ve j çalışmaları için,  $(\theta_i^*, v_i^*)$  ve  $(\theta_j^*, v_j^*)$  çiftlerinin sıralamaları aynı yönde farklıysa, yani i çalışması için  $\theta_i^*$  ve  $v_i^*$  sıralarının her ikisi de, j çalışması için ilgili sıralardan daha düşük veya her ikisi de daha yüksektir. Eğer iki değişkenin sıralarının karşılaştırılması ters yönde ise, sıraların uyumsuz olduğu, çeliştiği söylenir. Eğer x uyumlu sıra, y uyumsuz sıra varsa normalleştirilmiş test istatistiği z'dir (Rothstein ve ark., 2005; Jin, Zhou, & He, 2014).

$$z = \frac{x-y}{\sqrt{k(k-1)(2k+5)/18}} \quad (3.3)$$

Test istatistiği olan z, meta analizinde yanlılık yok sıfır hipotezi altında standart bir normal dağılıma göre asimptotik olarak dağılır. Meta analizinde yanlılık yok sıfır hipotezi  $\alpha$  anlamlılık düzeyinde  $|z| = z_{1-\alpha/2}$  ise reddedilir.

Standart hatalarının sıralaması  $\{s_i\}$ , varyansının  $\{v_i\}$  sıralamasına eşit olduğu için, Begg ve Mazumdar'ın testi, standart etki büyüklükleri ile standart hatalar arasındaki sıra korelasyonunun elde edilmesine eşdeğerdir (Rothstein ve ark., 2005). Bu test prosedürü, funnel grafiğinde görmeyi beklediğimiz çarpıklık kalıplarının  $\{\theta_i\}$  ve  $\{v_i\}$  arasında korelasyona yol açması gerektiği gerçeğine dayanarak, deneysel temellerde önerilmiştir (Begg, & Mazumdar, 1994; Jin ve ark., 2014). Yani yanlılığın olmadığı sıfır hipotezi altında, bir meta analizine dahil edilen çalışmalar arasında etki büyüklüğü ile kesinlik ( $\theta_i$  ve  $v_i$ ) arasında bir ilişki olmadığını varsayar.

### ***Egger'in Yöntemi***

Egger yöntemi ile asimetrinin değerlendirilmesi lineer regresyon yöntemine dayalı olarak yapılır.

$$E[z_i] = \beta_0 + \beta_1 prec_i \quad (3.4)$$

(3.4)'de gösterilen regresyon eşitliğinde  $E[z_i]$ ,  $\theta_i$ 'nin (i. çalışmadan elde edilen etki büyüklüğü tahmini) standart hataya bölüdüğü, standart etki büyüklüğüdür ( $z_i = \theta_i/s_i$ ).  $prec_i = 1/s_i$  olarak tanımlanmıştır.  $\beta_0$ , sabit ve  $\beta_1$ , eğimdir (Rothstein ve ark., 2005).

Egger testi, meta analizinde yanlılık yok sıfır hipotezi altında, k-2 serbestlik dereceli Student t dağılımına göre asimptotik olarak dağılmış sıfır olmayan bir  $\beta_0$  sabiti için test edilerek oluşturulmuştur. Meta analizinde yanlılık yok sıfır hipotezi  $\alpha$  anlamlılık düzeyinde  $|\hat{\beta}_0/s_i(\hat{\beta}_0)| = |t| > t_{k-2;1-\alpha/2}$  ise reddedilir. Test prosedürü, yanlılık varlığında doğrusallığın hala geçerli olduğu varsayımına dayanmaktadır.

Funnel grafiği simetrik ise,  $\beta_0 = 0$ 'dır, regresyon doğrusu orijinden geçer. Eğim ( $\beta_1$ ) ise, etki büyüklüğünü ve yönünü belirtir. Asimetri varsa, regresyon çizgisi orijinden geçmez,  $\beta_0$ , bir asimetri ölçüsü sağlar. Sıfırdan sapma ne kadar büyük olursa asimetri daha belirgin hale gelir (Egger ve ark., 1997; Rothstein ve ark., 2005).

Sabitin anlamlılığı meta analizi çalışmalarında genellikle  $\alpha=0,10$  anlamlılık düzeyine göre test edilir (Tang, & Liu, 2000). Sabitin negatif değerleri alması ise küçük çalışmaların, büyük çalışmalardan daha belirgin faydalı etkiler gösterdiğinin belirtir (Egger ve ark., 1997).

Egger ve arkadaşları (1997) etki tahmininin varyansının tersini kullanarak yani  $1/v$  ile ağırlıklandırarak bu testin alternatif bir versiyonunu daha önermiştir (Jin ve ark., 2014; Rothstein ve ark., 2005). z'nin "prec" üzerindeki ağırlıksız regresyonunun, ağırlıkların ( $1/v_i$ ) her çalışmada tahmini etki büyüklüğünün varyansı ile ters orantılı olduğu standart hatalar (s) üzerindeki etki büyüklüğünün ( $\theta$ ) ağırlıklı regresyonu ( $E[\theta_i] = \beta_1 + \beta_0 s_i$ ) ile aynı olduğu gösterilebilir. Ağırlıksız regresyon ile karşılaştırıldığında  $\beta_0$  ve  $\beta_1$  rolleri değişir, yani ilk regresyondaki sabit ikinci regresyondaki eğime karşılık gelir. Böylece, funnel grafiği asimetrisi için regresyon testi, standart hataya karşı etki büyüklüğünün bir funnel grafiğindeki doğrusal eğilim testine karşılık gelir (Rothstein ve ark., 2005).

### ***Schwarzer'in Yöntemi***

Schwarzer ve arkadaşları (2007) seyrek ikili sonuçları olan bir meta analizinde yanlılığı saptamak için Begg'in sıra korelasyon testinden türetilmiş bir test önerdi. Schwarzer ve arkadaşları, standartlaştırılmış hücre sayılarını ve yeni tanımlanan varyansı sırasıyla orijinal standart etki büyüklüğü ve varyansının yerine koymak için kullandı. Gözlenen hücre sayısının ( $a_i$ ) olasılığının, eşitlik (3.5)'de belirtilen merkezi olmayan hipergeometrik bir dağılımı takip ettiği varsayılmıştır.

$$P(a_i; \theta_i) = \frac{\binom{a_i+b_i}{a_i} \binom{c_i+d_i}{c_i} \theta_i^{a_i}}{\sum_{x \in U_i} \binom{a_i+b_i}{x} \binom{c_i+d_i}{a_i+c_i-x} \theta_i^x} \quad (3.5)$$

Burada,  $U_i$  sabit marjinal toplam verildiğinde  $a_i$  için olası değerler aralığıdır. Verilen bir  $\theta_i$  için merkezi olmayan hipergeometrik dağılımın ortalaması ve varyansı,

$$E(a_i; \theta_i) = \sum_{x \in U_i} x P(x; \theta_i) \quad (3.6)$$

$$Var(a_i; \theta_i) = \sum_{x \in U_i} (x - E(a_i; \theta_i))^2 P(x; \theta_i) \quad (3.7)$$

olarak hesaplanabilir.

Schwarzer'ın yönteminde, Mantel–Haenszel OR ( $\hat{\theta}_i^{MH}$ ) değerleri, bilinmeyen OR ( $\theta_i$ ) değerlerini tahmin için kullanılmıştır. Odds oranının koşullu maksimum olabilirlik tahmin edicisinin özellikleri ile,  $\hat{\theta}_i^{MH}$  için asimptotik varyans,  $1 / Var(a_i; \hat{\theta}_i^{MH})$  ile tahmin edilebilir. Schwarzer ve arkadaşları (2007) tarafından, Begg'in testine benzer şekilde Kendall  $\tau$ , funnel grafiğindeki asimetriyi iki şekilde test etmek için kullanılmıştır. İlk yaklaşım asimptotik varyansın ( $\hat{v}_i$ ) standartlaştırılmış etki büyüklüğünün hesaplanması prosedürü de dahil olmak üzere  $1 / Var(a_i; \hat{\theta}_i^{MH})$  ile değiştirilmesidir. Alternatif yaklaşım ise, standartlaştırılmış hücre sayısı ile  $1 / Var(a_i; \hat{\theta}_i^{MH})$  arasındaki korelasyonu test etmektir. Standartlaştırılmış hücre sayısı  $s(a_i; \hat{\theta}_i^{MH}) = (a_i - E(a_i; \hat{\theta}_i^{MH})) / (Var(a_i; \hat{\theta}_i^{MH})^{1/2})$  olarak tanımlanmıştır (Jin ve ark., 2014; Schwarzer, Antes, & Schumacher, 2007).

### ***Arcsin Farkına Dayanan testler (Rucker'in Testleri)***

Rucker ve arkadaşları tarafından geliştirilen (2008), arcsin dönüşümlü etki büyüklüğüne ve binom rasgele değişkenlerin stabilize varyansına dayanan modifiye bir yöntemdir. Arcsin regresyon testi etki büyüklüğü olarak arcsin farkı olan  $\Delta_i$  seçilerek tanımlanır. Arcsin dönüşümlü gözlenen risk farkı ve yaklaşık varyans sırasıyla  $\Delta_i = \arcsin(a_i/(e_i))^{1/2} - \arcsin(b_i/(f_i))^{1/2}$  ve  $\Gamma_i = 1/4e_i + 1/4(f_i)$  olarak tanımlandığında, orijinal  $\theta_i^*$  ve  $\hat{v}_i$  yerine  $\Delta_i$  ve  $\Gamma_i$  değiştirilerek Begg, Egger ve Thompson testlerinin özel bir formu olarak gerçekleştirilebilir (Jin ve ark., 2014). Varyans tahmini yalnızca satır toplamlarına, yani örneklem hacimlerine bağlıdır, etki büyüklüğü tahminine bağlı değildir.



### ***Macaskill'in Yöntemi***

Macaskill ve arkadaşları (2001), tahmin edilen etki büyüklüğü ( $\hat{\theta}_i$ ) bağımlı değişken ve çalışma büyüklüğü ( $n_i$ ) bağımsız değişken olarak kullanılan ağırlıklı bir regresyon yaklaşımı kabul etmişlerdir. Gözlemler, tahminin ters varyansına göre (FIV) veya her çalışma için birleştirilmiş varyansın tersine göre (FPV) ağırlıklandırılır. Ağırlık  $1/\hat{v}_i$  veya  $[1/(a_i + c_i) + 1/(b_i + d_i)]^{-1}$  olduğunda ağırlıklı regresyon modeli  $\hat{\theta}_i = \beta_0 + \beta_1 n_i + \varepsilon_i$  olarak verilir. Herhangi bir yanlılık olmadığında, regresyon eğimi olan  $\beta_1$  beklenen bir sıfır değerine sahiptir. Sıfır olmayan bir eğim, muhtemel yanlılık nedeniyle, etki büyüklüğü ile örneklem hacmi arasında bir ilişki olduğunu gösterir (Jin ve ark., 2014; Macaskill, Walter, & Irwig, 2001).

### ***Harbord'un Yöntemi***

Harbord ve arkadaşları (2006), her denemenin etki büyüklüğünün tahmini için etkin skor ( $Z_i = a_i - (a_i + c_i) \times e_i/n_i$ ) ve varyans skoruna ( $V_i = (a_i + c_i) \times (b_i + d_i) \times e_i \times f_i/[n_i^2 \times (n_i - 1)]$ ) dayanan ikili veriler için Egger testinin modifikasyonu olan bir test önerdiler. Bu tahminler ve yaklaşık standart hataları daha sonra, Egger ve arkadaşları (1997) tarafından önerilmiş olan regresyon testindeki etki büyüklüğü ve standart hatalar yerine modifiye edilmiş regresyon testinde kullanılır.  $V$  ağırlığı ile  $Z/V$  ve  $1/\sqrt{V}$ 'nin ağırlıklı regresyon modeli  $Z_i/V_i = \beta_0 + \beta_1 1/\sqrt{V_i} + \varepsilon_i$ 'dir. Burada  $V_i$  ağırlık,  $\varepsilon_i \sim N(0, \hat{v}_i/V_i \times \phi)$ .  $\beta_1 = 0$  olan sıfır hipotezi yanlılığın yokluğu ile ilgilidir. İki yönlü t testi ile test edilir (Harbord, Harris, & Sterne, 2009; Jin ve ark., 2014). Bu testin avantajı, varyans tahmininin sadece marjinal toplamlara bağlı olmasıdır (Schwarzer ve ark., 2015).

### ***Peters'in Yöntemi***

Peters ve arkadaşları (2006), temelinde toplam örneklem hacminin tersini bağımsız değişken olarak kullanan Macaskill'in funnel regresyon yöntemi (FPV) olan modifiye edilmiş bir regresyon yaklaşımı önermiştir. Ağırlığın  $[1/(a_i + c_i) + 1/(b_i + d_i)]^{-1}$  ve  $\varepsilon_i \sim N(0, se_i^2 \times \phi)$  olduğu ağırlıklı regresyon modeli  $\hat{\theta}_i = \beta_0 + \beta_1 1/(a_i + b_i + c_i + d_i) + \varepsilon_i$  'dir.  $\beta_1 = 0$  olan sıfır hipotezi yanlılığın yokluğu ile ilgilidir. Regresyon çizgisinin eğimi, sıfır hipotezinin sıfır eğim olduğu iki taraflı t testi ile test edilir (Harbord ve ark., 2009; Jin ve ark., 2014).

### ***Thompson'ın Yöntemi***

Egger testinin çalışmalar arası heterojenliğe izin veren bir çeşidi olarak Thompson ve Sharp (1999) tarafından önerilmiştir. Test istatistiği, toplanabilir çalışmalar arası varyans bileşeni için moment tahmin yöntemi kullanılarak, standart hata üzerindeki etki büyüklüğünün ağırlıklı bir doğrusal regresyonuna dayanmaktadır. Test istatistiği 2 serbestlik dereceli bir t dağılımına uyar. Çalışmalar arası varyans  $\tau^2$  olarak verildiğinde ve  $w_i = 1/(s_i^2 + \tau^2)$  ağırlıklı regresyon modeli  $\varepsilon_i \sim N(0, s_i^2 + \tau^2)$  dağılımı ile  $\theta_i = \beta_0 + \beta_1 \times s_i + \varepsilon_i$  dir.  $\beta_0$  ve  $\beta_1$ 'nin maksimum olasılık tahminleri,  $w_i$  ağırlığı ile en küçük kareler regresyonuyla elde edilir. Sıfır hipotezi yanlılık olmadığına karşılık gelen  $\beta_1 = 0$ 'dır.

### 3. GEREÇ VE YÖNTEM

Çalışmada, literatürde yer alan ve ikili değer alan verilerin meta analizinde funnel grafiği asimetri değerlendirmesi için kullanılan parametrik ve parametrik olmayan yanlılık belirleme testleri arasında bulunan Begg'in, Egger'in, Schwarzer'in, Harbord'in, Thompson'un testlerinin çeşitli simülasyon senaryoları altında Tip-I hata oranları ve güçleri yönünden performanslarının karşılaştırılması amaçlanmıştır. Tez çalışmasında farklı hastalık-sonuç oranlarında 1.000.000 birimden oluşan varsayımsal anakütller oluşturulmuştur. Testlerin performansları, Cochran Q testine ( $\alpha=0,10$ ) göre meta analizine alınan çalışmaların homojen olması sağlanacak şekilde yansızlık ve farklı yanlılık derecelerine göre, farklı çalışma sayıları (k), farklı örneklem hacimleri (n) ve farklı hastalık-sonuç oranları (P) durumlarını içeren simülasyon senaryoları sonucunda değerlendirilmiştir. Meta analizinde yanlılık olmadığı duruma ait simülasyon senaryolarının her biri için Tip-I hata oranları, 1000 tekrar sonunda reddedilen  $H_0$  hipotezlerinin sayıları belirlendikten sonra elde edilmiştir. Meta analizinde farklı yanlılık derecelerinde yanlılık oluşturulduğu durumlara ait simülasyon senaryolarının her biri için güç, 1000 tekrar sonunda reddedilen  $H_0$  hipotezlerinin sayıları belirlendikten sonra elde edilmiştir.

Simülasyon çalışmalarında meta analizinde yanlılık belirlemede kullanılan testler için Tip-I hata olasılığı  $\alpha=0,10$  olarak alınmıştır. Simülasyon sonuçları sonrasında elde edilen Tip-I hata olasılığının yorumlanması amacıyla literatürde Peterson (2002) tarafından önerilen kriterler çalışmamızda dikkate alınmıştır. Peterson (2002) kriterlerine göre, denemeler sonucunda üretilen Tip-I hata olasılığı,  $\alpha=0,10$  olması durumunda  $0,09-0,11$  ( $\alpha \pm 0,1\alpha$ ) arasında ise sağlam olarak,  $0,075-0,125$  ( $\alpha \pm 0,25\alpha$ ) arasında ise orta derecede sağlam olarak belirlenmiştir. İki durumda da Tip-I hata olasılıklarını üreten testlerin Tip-I hatayı koruma yönündeki performansları yeterli olarak nitelendirilebilir. Ayrıca; testlerinin Tip-I hatayı koruma kriterine göre gösterdikleri performanslar incelendiğinde Tip-I hata oranını nominal düzeyinin (0,10) altında tahmin etme eğiliminde olduklarında konservatif bir tutum, Tip-I hata oranını

nominal düzeyinin (0,10) üstünde tahmin etme eğiliminde olduklarında ise liberal bir tutum gösterdikleri söylenebilir (Hsiung, & Olejnik, 1996).

### 3.1. Simülasyon Senaryoları

Simülasyon çalışmasında, binom dağılımı ile kontrol grubu ( $H^-$ ) için sonuç pozitif olması ( $S^+$ ) olasılığı  $P(S^+|H^-)=0,50$  olan  $N_K=1.000.000$  ve hasta grubu ( $H^+$ ) için sonuç pozitif olması ( $S^+$ ) olasılıkları ise  $P(S^+|H^+)={0,50; 0,60; 0,70; 0,80; 0,90}$  olan  $N_H=1.000.000$  olan varsayımsal anakütller oluşturulmuştur.

Üretilen her bir varsayımsal anakütleden örneklem hacimleri  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  rasgele örneklem ve meta analizine alınan çalışma sayıları ise  $k= 8, 10, 15, 20, 25, 50, 100$  ve  $150$  olacak şekilde ayarlanmıştır.

İlk bölümde, yanlılığın olmadığı, ikinci bölümde ise funnel grafiğinde asimetri oluşturmak amacıyla grafiğin sağ tarafında yer alan ve meta analizine dahil edilmiş olan çalışmaların sıra ile %10, %20, %40 ve %60'ı her bir veri kümesinden çıkartılarak yanlılık oluşturulduğu durumlara ait farklı örneklem hacimleri, farklı çalışma sayıları ve farklı sonuç pozitif olasılıklarına sahip verilerden OR'ler üretilerek yanlılık belirleme yöntemleri uygulanmıştır. Simülasyon çalışmasında tekrar sayısı 1000 olarak alınmıştır. Simülasyon senaryolarının daha iyi anlaşılması için ilgili karakteristikler Tablo 2'de özetlenmiştir. Simülasyon çalışması R-project 3.6.0 açık kaynak kodlu program kullanılarak yapılmıştır. Meta analizleri için R-project programındaki readr (Wickham, & Hester, 2022), dplyr (Wickham, François, & Henry 2022), metafor (Viechtbauer, 2010), metasens (Schwarzer, Carpenter, & Rücker 2022), meta (Balduzzi, Rücker, & Schwarzer, 2019), epiR (Stevenson, & Sergeant, 2022) paketleri ve "metabin" fonksiyonu kullanılmıştır.

**Tablo 2.** Simülasyon senaryolarında yanlılığın olmadığı, % 10, %20, %40 ve %60 yanlılık durumlarında kullanılan değişkenler

Kontrol grubu için sonuç pozitif olasılığı	Hasta grupları için sonuç pozitif olasılıkları	Kontrol ve hasta grubundan rastgele çekilecek örneklem sayıları	Meta analizine alınan çalışma sayısı
$P(S^+ H^-)=0,50$ olarak ayarlandı.	$P(S^+ H^+)=0,50; 0,60; 0,70; 0,80; 0,90$ olacak şekilde ayarlandı.	$n_K=n_H=8, 10, 12, 15, 20, 25, 50$ olacak şekilde ayarlandı	$k=8, 10, 15, 20, 25, 50, 100$ ve $150$ olacak şekilde ayarlandı
$N_K=1.000.000$	Belirlenen her bir olasılık değeri için $N_H=1.000.000$ olacak şekilde		

## 4. BULGULAR

### 4.1. Meta Analizde Yanlılık Olmadığı Duruma Ait Bulgular

Meta analizinde yanlılık olmadığı durum dikkate alınarak yapılan simülasyon çalışması ile elde edilen veriler analiz edilmiş ve analizlerden elde edilen bulgular tablolar (Tablo 3 –7) aracılığıyla sunulmuştur. Hasta grubunda pozitif sonuç oranlarına göre ayrı ayrı sunulan tablolarda anakütle OR değerleri,  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  olan örneklem hacimleri ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  olan meta analizine alınan çalışma sayılarına göre funnel grafiği asimetrisinin değerlendirilmesi ile meta analizinde yanlılık olup olmadığına karar vermemizi sağlayan Begg, Egger, Schwarzer, Harbord ve Thompson testlerine ait Tip-I hata oranları verilmiştir. Nominal değer yani Tip-I hata düzeyi  $\alpha=0,10$  olarak alınmıştır.

Yanlılığın olmadığı,  $OR=1,00$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,50$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin Tip-I hata yönünden performanslarına ait simülasyon sonuçları Tablo 3'te gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Schwarzer testine yakın bir performansla Harbord testi ve daha sonra sadece meta analizine alınan çalışma sayısı  $k=8$  olduğu durumda Tip-I hata düzeyini koruma eğiliminde olan Thompson testi takip etmektedir. Tablo 3'te yer alan testler arasında Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg ve Egger testleridir. Bu iki testin benzer sonuçlar verdiği ve her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir. Schwarzer ve Harbord testlerinin Tip-I hata oranları birbirine ve

belirlenen nominal değere yakındır. Meta analizine alınan çalışma sayısı arttığında nominal değerden Schwarzer testinin az da olsa artış yönünde, Harbord testinin ise azalış yönünde saptığı görülmektedir. Örneklem hacminin az olduğu durumlarda Harbord testinin nominal değerden azalış, Schwarzer testinin az da olsa nominal değerden artış yönünde saptığı görülmektedir. Schwarzer testi meta analizine alınan çalışma sayısı  $\leq 25$  ve örneklemdeki birim sayısı  $n \leq 10$  gibi az olduğunda Tip-I hatayı koruma yönünde iyi performans göstermektedir. Thompson testi, daha çok meta analizine alınan çalışma sayısı az ise ( $k=8$ ) Tip-I hata olasılık düzeyini koruyabilmektedir. Meta analizine alınan çalışma sayısı arttıkça Thompson testinin Tip-I hata oranları artarak nominal değerden sapmakta ve Egger testi Tip-I hata oranlarına benzer sonuçlar vermektedir.

**Tablo 3.** Yanlılık olmadığı durumda  $OR=1,00$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,50$  için Tip-I hata oranlarına ait simülasyon sonuçları

<b>k</b>	<b><math>n_H=n_K</math></b>	<b>Begg Yöntemi</b>	<b>Egger Yöntemi</b>	<b>Schwarzer Yöntemi</b>	<b>Harbord Yöntemi</b>	<b>Thompson Yöntemi</b>
<b>8</b>	<b>8</b>	0,184	0,248	0,106	0,084	0,089
	<b>10</b>	0,209	0,322	0,097	0,072	0,115
	<b>12</b>	0,186	0,266	0,121	0,100	0,115
	<b>15</b>	0,217	0,315	0,115	0,112	0,130
	<b>20</b>	0,168	0,282	0,102	0,089	0,131
	<b>25</b>	0,214	0,310	0,113	0,102	0,130
<b>10</b>	<b>8</b>	0,160	0,269	0,108	0,108	0,121
	<b>10</b>	0,206	0,275	0,113	0,064	0,148
	<b>12</b>	0,246	0,330	0,112	0,089	0,203
	<b>15</b>	0,182	0,291	0,128	0,096	0,170
	<b>20</b>	0,231	0,340	0,111	0,115	0,195
	<b>25</b>	0,182	0,271	0,103	0,095	0,158
<b>15</b>	<b>8</b>	0,203	0,316	0,103	0,091	0,173
	<b>10</b>	0,193	0,304	0,105	0,108	0,173
	<b>12</b>	0,209	0,288	0,104	0,061	0,200
	<b>15</b>	0,231	0,324	0,105	0,081	0,244
	<b>20</b>	0,209	0,324	0,112	0,090	0,242
	<b>25</b>	0,207	0,306	0,137	0,124	0,237
<b>20</b>	<b>8</b>	0,206	0,301	0,103	0,089	0,228
	<b>10</b>	0,205	0,314	0,111	0,104	0,242
	<b>12</b>	0,214	0,317	0,103	0,101	0,211
	<b>15</b>	0,221	0,307	0,125	0,061	0,242
	<b>20</b>	0,222	0,307	0,096	0,070	0,252
	<b>25</b>	0,209	0,327	0,136	0,093	0,260
<b>25</b>	<b>8</b>	0,226	0,312	0,117	0,090	0,256
	<b>10</b>	0,211	0,324	0,113	0,113	0,257
	<b>12</b>	0,199	0,346	0,095	0,100	0,279
	<b>15</b>	0,197	0,309	0,103	0,102	0,251
	<b>20</b>	0,234	0,279	0,125	0,056	0,238
	<b>25</b>	0,216	0,348	0,125	0,088	0,284
<b>50</b>	<b>8</b>	0,221	0,318	0,106	0,091	0,273
	<b>10</b>	0,230	0,344	0,110	0,106	0,295
	<b>12</b>	0,200	0,322	0,110	0,103	0,270
	<b>15</b>	0,208	0,338	0,106	0,103	0,291
	<b>20</b>	0,184	0,322	0,109	0,098	0,265
	<b>25</b>	0,279	0,273	0,129	0,050	0,233
<b>100</b>	<b>8</b>	0,255	0,308	0,128	0,066	0,289
	<b>10</b>	0,228	0,330	0,121	0,084	0,305
	<b>12</b>	0,240	0,326	0,110	0,099	0,296
	<b>15</b>	0,218	0,325	0,107	0,096	0,298
	<b>20</b>	0,217	0,336	0,098	0,100	0,314
	<b>25</b>	0,210	0,328	0,088	0,087	0,303
<b>150</b>	<b>8</b>	0,302	0,252	0,113	0,043	0,214
	<b>10</b>	0,277	0,330	0,120	0,050	0,304
	<b>12</b>	0,291	0,338	0,130	0,086	0,320
	<b>15</b>	0,276	0,347	0,123	0,107	0,329
	<b>20</b>	0,230	0,364	0,114	0,100	0,336
	<b>25</b>	0,227	0,349	0,100	0,106	0,324
<b>150</b>	<b>8</b>	0,214	0,344	0,115	0,097	0,326
	<b>10</b>	0,325	0,265	0,132	0,048	0,234
	<b>12</b>	0,297	0,318	0,128	0,064	0,306
	<b>15</b>	0,297	0,332	0,131	0,081	0,319
	<b>20</b>	0,282	0,340	0,107	0,097	0,327
	<b>25</b>	0,250	0,344	0,126	0,102	0,330
<b>150</b>	<b>25</b>	0,247	0,356	0,108	0,101	0,343
	<b>50</b>	0,220	0,338	0,095	0,093	0,322

Yanlılığın olmadığı,  $OR=1,50$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,60$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin Tip-I hata yönünden performanslarına ait simülasyon sonuçları Tablo 4'te gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha= 0,10$ ) koruyamamasıyla birlikte, en iyi performansı Harbord testi göstermiştir. Harbord testini,  $k \leq 25$  olduğu her durumda Tip-I hatayı kontrol altında tutan Schwarzer testi takip etmektedir. Tablo 4'te yer alan testler arasında Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler birbirine yakın Tip-I hata oranları ile Begg, Egger ve Thompson testleridir. Harbord testi  $k \geq 100$  olduğu her durumda Tip-I hatayı kontrol altında tutma eğilimindedir. Schwarzer testinin meta analizine alınan çalışma sayısı arttıkça Tip-I hata düzeyini artış yönünde, Harbord testinin ise meta analizine alınan çalışma sayısı azaldıkça Tip-I hata düzeyini azalış yönünde sapmalı olarak tahmin etme eğilimi artmıştır. Schwarzer testinin, meta analizine alınan çalışma sayısı  $k=50,100$  ve örneklemdeki birim sayısı azken Tip-I hata düzeyini koruma yönünde eğilimi vardır. Meta analizine alınan çalışma ve örneklemdeki birim sayısının artması ile Begg, Egger ve Thompson testlerinin Tip-I hata oranlarında nominal düzeyden sapmanın da artış yönünde arttığı gözlenmektedir. OR değerinin 1,00'ten 1,50'e artmasıyla Begg, Egger, Thompson testlerinde Tip-I hata oranlarının ve dolayısıyla nominal değerden artış yönünde sapmanın arttığı gözlenmiştir. OR değerinin 1,00'ten 1,50'e artmasıyla Schwarzer testinin Tip-I hata düzeyini koruma yönündeki performansı düşerken, Harbord testinin ise Tip-I hata düzeyini koruma yönündeki performansı yükselmiştir.



**Tablo 4.** Yanlılık olmadığı durumda  $OR=1,50$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,60$  için Tip-I hata oranlarına ait simülasyon sonuçları

<b>k</b>	<b><math>n_H=n_K</math></b>	<b>Begg Yöntemi</b>	<b>Egger Yöntemi</b>	<b>Schwarzer Yöntemi</b>	<b>Harbord Yöntemi</b>	<b>Thompson Yöntemi</b>
<b>8</b>	<b>8</b>	0,319	0,334	0,101	0,069	0,134
	<b>10</b>	0,354	0,385	0,092	0,082	0,185
	<b>12</b>	0,402	0,426	0,113	0,097	0,190
	<b>15</b>	0,428	0,420	0,095	0,095	0,212
	<b>20</b>	0,479	0,482	0,119	0,106	0,237
	<b>25</b>	0,498	0,482	0,095	0,104	0,227
	<b>50</b>	0,588	0,590	0,107	0,105	0,314
<b>10</b>	<b>8</b>	0,368	0,387	0,107	0,069	0,197
	<b>10</b>	0,430	0,460	0,098	0,090	0,291
	<b>12</b>	0,456	0,455	0,117	0,092	0,273
	<b>15</b>	0,514	0,477	0,095	0,087	0,308
	<b>20</b>	0,548	0,537	0,103	0,102	0,368
	<b>25</b>	0,593	0,549	0,110	0,105	0,348
	<b>50</b>	0,679	0,647	0,102	0,104	0,456
<b>15</b>	<b>8</b>	0,470	0,435	0,097	0,083	0,320
	<b>10</b>	0,567	0,508	0,087	0,070	0,401
	<b>12</b>	0,598	0,543	0,112	0,087	0,434
	<b>15</b>	0,676	0,591	0,106	0,104	0,478
	<b>20</b>	0,734	0,655	0,109	0,090	0,545
	<b>25</b>	0,790	0,669	0,108	0,103	0,565
	<b>50</b>	0,864	0,760	0,115	0,096	0,660
<b>20</b>	<b>8</b>	0,566	0,473	0,104	0,082	0,388
	<b>10</b>	0,634	0,552	0,106	0,078	0,480
	<b>12</b>	0,691	0,602	0,109	0,093	0,522
	<b>15</b>	0,803	0,675	0,110	0,114	0,602
	<b>20</b>	0,846	0,714	0,097	0,090	0,640
	<b>25</b>	0,892	0,751	0,106	0,093	0,665
	<b>50</b>	0,950	0,860	0,110	0,091	0,798
<b>25</b>	<b>8</b>	0,633	0,505	0,104	0,068	0,441
	<b>10</b>	0,731	0,613	0,102	0,077	0,557
	<b>12</b>	0,771	0,658	0,115	0,094	0,602
	<b>15</b>	0,870	0,714	0,125	0,109	0,662
	<b>20</b>	0,910	0,749	0,105	0,086	0,704
	<b>25</b>	0,931	0,817	0,124	0,121	0,751
	<b>50</b>	0,983	0,910	0,102	0,102	0,872
<b>50</b>	<b>8</b>	0,891	0,694	0,112	0,075	0,651
	<b>10</b>	0,942	0,778	0,116	0,090	0,748
	<b>12</b>	0,948	0,825	0,109	0,069	0,792
	<b>15</b>	0,985	0,872	0,130	0,104	0,858
	<b>20</b>	0,996	0,906	0,127	0,076	0,892
	<b>25</b>	0,997	0,953	0,151	0,108	0,941
	<b>50</b>	1,000	0,987	0,107	0,090	0,985
<b>100</b>	<b>8</b>	0,992	0,890	0,110	0,075	0,864
	<b>10</b>	0,997	0,922	0,120	0,081	0,907
	<b>12</b>	0,998	0,949	0,150	0,075	0,943
	<b>15</b>	1,000	0,973	0,189	0,116	0,967
	<b>20</b>	1,000	0,991	0,165	0,090	0,991
	<b>25</b>	1,000	0,993	0,161	0,104	0,992
	<b>50</b>	1,000	1,000	0,147	0,083	1,000
<b>150</b>	<b>8</b>	1,000	0,955	0,107	0,092	0,948
	<b>10</b>	0,999	0,975	0,134	0,094	0,974
	<b>12</b>	1,000	0,988	0,193	0,083	0,986
	<b>15</b>	1,000	0,997	0,205	0,120	0,996
	<b>20</b>	1,000	1,000	0,167	0,083	1,000
	<b>25</b>	1,000	1,000	0,219	0,102	1,000
	<b>50</b>	1,000	1,000	0,158	0,104	1,000

Yanlılığın olmadığı,  $OR=2,34$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,70$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin Tip-I hata yönünden performanslarına ait simülasyon sonuçları Tablo 5'te gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Schwarzer testini, Schwarzer testine yakın bir performansla Harbord testi takip etmektedir. Tablo 5'te yer alan testler arasında Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg, Egger ve Thompson testleridir. Meta analizine alınan çalışma ve örneklemdeki birim sayısının artması ile Tip-I hata düzeyini her koşulda artış yönünde sapmalı olarak tahmin eden Begg, Egger ve Thompson testlerinin Tip-I hata oranlarında artış olduğu gözlenmektedir. Meta analizine alınan çalışma sayısı fazla olduğunda ( $k \geq 50$ ) örnekleme alınan birim sayısı azaldıkça Harbord testi Tip-I hata oranları artmış ve nominal değerden uzaklaşarak Tip-I hata düzeyini sapmalı olarak tahmin etmeye başlamıştır. OR değerinin 1,50'ten 2,34'e artmasıyla Begg, Egger ve Thompson testlerinde Tip-I hata oranlarının ve dolayısıyla nominal değerden sapmanın arttığı gözlenmiştir. Schwarzer testinin, OR değeri 1,00'ten 1,50'e çıktığında düşen Tip-I hata düzeyini koruma yönündeki performansının, OR değeri 2,34 olduğu simülasyon sonuçlarında az da olsa yükseldiği gözlenmektedir. OR değerinin 1,50'ten 2,34'e artması ile Harbord testinin Tip-I hata düzeyini koruma yönündeki performansı az da olsa azalmaya başlamıştır.

**Tablo 5.** Yanlılık olmadığı durumda  $OR=2,34$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,70$  için Tip-I hata oranlarına ait simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,502	0,496	0,112	0,086	0,211
	10	0,573	0,559	0,104	0,087	0,273
	12	0,561	0,580	0,102	0,097	0,259
	15	0,612	0,636	0,094	0,086	0,318
	20	0,621	0,639	0,106	0,099	0,328
	25	0,649	0,689	0,108	0,099	0,368
10	50	0,646	0,716	0,087	0,099	0,411
	8	0,587	0,557	0,097	0,077	0,307
	10	0,683	0,637	0,085	0,084	0,413
	12	0,694	0,666	0,094	0,088	0,430
	15	0,742	0,733	0,082	0,075	0,479
	20	0,750	0,726	0,096	0,104	0,503
15	25	0,749	0,757	0,100	0,109	0,530
	50	0,754	0,792	0,098	0,090	0,588
	8	0,766	0,617	0,092	0,088	0,486
	10	0,868	0,756	0,074	0,086	0,625
	12	0,864	0,777	0,072	0,083	0,669
	15	0,903	0,832	0,093	0,074	0,737
20	20	0,912	0,856	0,112	0,087	0,750
	25	0,928	0,897	0,108	0,113	0,795
	50	0,915	0,934	0,098	0,097	0,844
	8	0,862	0,723	0,089	0,094	0,612
	10	0,941	0,817	0,080	0,081	0,748
	12	0,949	0,860	0,091	0,094	0,787
25	15	0,966	0,904	0,103	0,093	0,841
	20	0,976	0,935	0,112	0,089	0,876
	25	0,979	0,958	0,097	0,098	0,893
	50	0,978	0,978	0,100	0,090	0,946
	8	0,947	0,793	0,100	0,104	0,697
	10	0,968	0,875	0,097	0,079	0,821
50	12	0,981	0,913	0,093	0,095	0,868
	15	0,990	0,942	0,097	0,082	0,907
	20	0,994	0,971	0,094	0,087	0,941
	25	0,994	0,970	0,105	0,104	0,948
	50	0,994	0,992	0,099	0,089	0,978
	8	0,997	0,967	0,098	0,150	0,946
100	10	1,000	0,985	0,096	0,110	0,978
	12	1,000	0,991	0,099	0,095	0,989
	15	1,000	0,995	0,106	0,084	0,992
	20	1,000	1,000	0,107	0,085	0,998
	25	1,000	1,000	0,116	0,107	1,000
	50	1,000	1,000	0,098	0,102	1,000
150	8	1,000	1,000	0,124	0,285	1,000
	10	1,000	1,000	0,088	0,183	1,000
	12	1,000	1,000	0,095	0,112	1,000
	15	1,000	1,000	0,129	0,115	1,000
	20	1,000	1,000	0,123	0,095	1,000
	25	1,000	1,000	0,118	0,091	1,000
150	50	1,000	1,000	0,099	0,093	1,000
	8	1,000	1,000	0,127	0,423	1,000
	10	1,000	1,000	0,090	0,232	1,000
	12	1,000	1,000	0,094	0,137	1,000
	15	1,000	1,000	0,104	0,097	1,000
	20	1,000	1,000	0,119	0,091	1,000
150	25	1,000	1,000	0,132	0,097	1,000
	50	1,000	1,000	0,092	0,094	1,000

Yanlılığın olmadığı,  $OR=4,02$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,80$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin Tip-I hata yönünden performanslarına ait simülasyon sonuçları Tablo 6'da gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Schwarzer testini, Schwarzer testine yakın bir performansla Harbord testi takip etmektedir. Tablo 6'da yer alan testler arasında Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg, Egger ve Thompson testleridir. Meta analizine alınan çalışma sayısı küçük olduğunda (8, 10, 15 gibi) Schwarzer ve Harbord testlerinin Tip-I hata düzeyini koruma yönündeki performansları daha iyidir. Meta analizine alınan çalışma sayısı arttığında (50,100,150 gibi) ve örneklem hacmi azaldığında Schwarzer ve Harbord testlerinin Tip-I hata oranları nominal değerden oldukça uzaklaşmakta ve bu testler Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etmeye başlamaktadır. Meta analizine alınan çalışma sayısı ve örneklem hacminin artması ile Tip-I hata olasılığını sapmalı olarak tahmin eden Begg, Egger ve Thompson testlerinin Tip-I hata oranlarında nominal değerden artış yönünde sapmanın arttığı gözlenmektedir. OR değerinin 2,34'den 4,02'e artmasıyla Begg, Egger ve Thompson testlerinin Tip-I hata oranlarında artış görülmüş ve dolayısıyla nominal değerden sapma artmıştır. OR değerinin 2,34'den 4,02'e artmasıyla Schwarzer ve Harbord testlerinin performansı düşmüş ve bu testler OR değerinin 1,00 olduğu durumdan bu yana en kötü performansı göstermişlerdir.

**Tablo 6.** Yanlılık olmadığı durumda  $OR=4,02$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,80$  için Tip-I hata oranlarına ait simülasyon sonuçları

<b>k</b>	<b><math>n_H=n_K</math></b>	<b>Begg Yöntemi</b>	<b>Egger Yöntemi</b>	<b>Schwarzer Yöntemi</b>	<b>Harbord Yöntemi</b>	<b>Thompson Yöntemi</b>
<b>8</b>	<b>8</b>	0,438	0,563	0,149	0,153	0,178
	<b>10</b>	0,555	0,613	0,117	0,118	0,252
	<b>12</b>	0,577	0,654	0,111	0,100	0,240
	<b>15</b>	0,625	0,729	0,108	0,088	0,355
	<b>20</b>	0,664	0,776	0,099	0,085	0,402
	<b>25</b>	0,691	0,794	0,082	0,071	0,418
<b>10</b>	<b>8</b>	0,550	0,583	0,179	0,177	0,286
	<b>10</b>	0,647	0,716	0,126	0,114	0,404
	<b>12</b>	0,696	0,723	0,090	0,096	0,424
	<b>15</b>	0,747	0,827	0,116	0,088	0,516
	<b>20</b>	0,795	0,862	0,093	0,090	0,603
	<b>25</b>	0,798	0,887	0,103	0,092	0,621
<b>15</b>	<b>8</b>	0,729	0,695	0,230	0,224	0,485
	<b>10</b>	0,854	0,852	0,138	0,155	0,683
	<b>12</b>	0,887	0,883	0,111	0,130	0,723
	<b>15</b>	0,937	0,942	0,105	0,091	0,815
	<b>20</b>	0,956	0,969	0,098	0,085	0,869
	<b>25</b>	0,949	0,974	0,078	0,082	0,881
<b>20</b>	<b>8</b>	0,835	0,820	0,286	0,289	0,651
	<b>10</b>	0,920	0,912	0,177	0,190	0,817
	<b>12</b>	0,946	0,936	0,125	0,154	0,872
	<b>15</b>	0,981	0,978	0,106	0,095	0,932
	<b>20</b>	0,988	0,993	0,110	0,086	0,959
	<b>25</b>	0,988	0,991	0,108	0,092	0,969
<b>25</b>	<b>8</b>	0,897	0,878	0,321	0,356	0,748
	<b>10</b>	0,968	0,944	0,203	0,230	0,888
	<b>12</b>	0,978	0,976	0,123	0,150	0,949
	<b>15</b>	0,997	0,988	0,121	0,111	0,970
	<b>20</b>	0,999	0,998	0,113	0,102	0,990
	<b>25</b>	0,999	0,998	0,092	0,075	0,992
<b>50</b>	<b>8</b>	0,998	1,000	0,098	0,101	0,993
	<b>8</b>	0,997	0,992	0,558	0,640	0,975
	<b>10</b>	1,000	0,999	0,327	0,431	0,997
	<b>12</b>	1,000	1,000	0,198	0,299	1,000
	<b>15</b>	1,000	1,000	0,120	0,162	1,000
	<b>20</b>	1,000	1,000	0,089	0,098	1,000
<b>100</b>	<b>25</b>	1,000	1,000	0,101	0,102	1,000
	<b>50</b>	1,000	1,000	0,105	0,103	1,000
	<b>8</b>	1,000	1,000	0,799	0,905	1,000
	<b>10</b>	1,000	1,000	0,551	0,732	1,000
	<b>12</b>	1,000	1,000	0,296	0,491	1,000
	<b>15</b>	1,000	1,000	0,159	0,275	1,000
<b>150</b>	<b>20</b>	1,000	1,000	0,093	0,131	1,000
	<b>25</b>	1,000	1,000	0,115	0,123	1,000
	<b>50</b>	1,000	1,000	0,084	0,094	1,000
	<b>8</b>	1,000	1,000	0,928	0,981	1,000
	<b>10</b>	1,000	1,000	0,701	0,872	1,000
	<b>12</b>	1,000	1,000	0,408	0,673	1,000
<b>150</b>	<b>15</b>	1,000	1,000	0,192	0,353	1,000
	<b>20</b>	1,000	1,000	0,104	0,155	1,000
	<b>25</b>	1,000	1,000	0,117	0,130	1,000
	<b>50</b>	1,000	1,000	0,094	0,102	1,000

Yanlılığın olmadığı,  $OR=9,01$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,90$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin Tip-I hata yönünden performanslarına ait simülasyon sonuçları Tablo 7’de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Schwarzer testini, Harbord testi ve sadece meta analizine alınan çalışma sayısı  $k=8$ , örneklem hacmi  $n=8$  olduğu durumda Tip-I hata düzeyini koruyabilen Thompson testi takip etmektedir. Tablo 7’de yer alan testler arasında Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg ve Egger testleridir. Egger testinin Tip-I hata oranları her örneklem hacmi ve meta analizine alınan çalışma sayısı için diğer dört testten daha yüksektir. Meta analizine alınan çalışma sayısının artması ve örneklem hacminin azalması ile Schwarzer ve Harbord testlerinin Tip-I hata oranları nominal değerden artış yönünde oldukça uzaklaşmakta ve Tip-I hata düzeyini sapmalı olarak tahmin etmeye başlamaktadır. Örneklem hacmi 50 ise her meta analizine alınan çalışma sayısında Schwarzer testinin Tip-I hata düzeyini koruma yönündeki performansı iyidir. Harbord testi ise meta analizine alınan çalışma sayısı  $k \leq 20$  ve örneklem hacmi  $n=50$  olduğu durumda Tip-I hata düzeyini koruyabilmiştir. Meta analizine alınan çalışma sayısı ve örneklem hacminin artması ile Tip-I hata olasılığını artış yönünde sapmalı olarak tahmin eden Begg, Egger ve Thompson testlerinin Tip-I hata oranlarında nominal değerden sapmanın arttığı gözlenmektedir. OR değerinin 4,02’den 9,01’e artmasıyla Schwarzer ve Harbord testlerinin performansı düşmüş ve OR değerinin 1,00 olduğu durumdan bu yana en kötü performansı göstermişlerdir. OR değerinin 4,02’den 9,01’e artmasıyla Begg, Egger ve Thompson testlerinin Tip-I hata oranlarında azalış görülmüş ve nominal değerden sapma azalmıştır.

**Tablo 7.** Yanlılık olmadığı durumda  $OR=9,01$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,90$  için Tip-I hata oranlarına ait simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,347	0,468	0,487	0,436	0,091
	10	0,400	0,557	0,330	0,346	0,134
	12	0,467	0,616	0,257	0,277	0,161
	15	0,535	0,745	0,159	0,186	0,221
	20	0,662	0,834	0,123	0,147	0,293
	25	0,737	0,883	0,088	0,122	0,380
	50	0,799	0,932	0,107	0,107	0,559
10	8	0,391	0,513	0,591	0,527	0,149
	10	0,505	0,630	0,420	0,392	0,250
	12	0,561	0,728	0,315	0,341	0,299
	15	0,644	0,839	0,193	0,230	0,372
	20	0,788	0,915	0,130	0,185	0,505
	25	0,849	0,950	0,095	0,133	0,606
	50	0,921	0,980	0,089	0,092	0,751
15	8	0,463	0,602	0,744	0,699	0,289
	10	0,643	0,771	0,600	0,569	0,457
	12	0,735	0,855	0,455	0,481	0,561
	15	0,814	0,952	0,295	0,335	0,684
	20	0,948	0,989	0,159	0,219	0,839
	25	0,967	0,989	0,097	0,149	0,900
	50	0,992	1,000	0,106	0,111	0,959
20	8	0,575	0,712	0,851	0,828	0,430
	10	0,739	0,866	0,700	0,710	0,600
	12	0,840	0,933	0,530	0,581	0,749
	15	0,909	0,982	0,365	0,416	0,871
	20	0,990	0,999	0,184	0,257	0,958
	25	0,992	1,000	0,120	0,195	0,976
	50	1,000	1,000	0,094	0,117	0,998
25	8	0,633	0,783	0,927	0,900	0,538
	10	0,810	0,926	0,789	0,799	0,745
	12	0,893	0,964	0,637	0,681	0,849
	15	0,960	0,994	0,435	0,495	0,960
	20	0,998	0,999	0,215	0,346	0,988
	25	0,999	0,999	0,136	0,241	0,988
	50	0,999	1,000	0,098	0,141	1,000
50	8	0,867	0,962	0,998	0,996	0,874
	10	0,969	0,997	0,970	0,971	0,978
	12	0,993	1,000	0,894	0,925	0,995
	15	1,000	1,000	0,746	0,813	1,000
	20	1,000	1,000	0,371	0,562	1,000
	25	1,000	1,000	0,207	0,421	1,000
	50	1,000	1,000	0,100	0,193	1,000
100	8	0,980	1,000	1,000	1,000	0,992
	10	0,998	1,000	1,000	1,000	1,000
	12	1,000	1,000	0,990	0,998	1,000
	15	1,000	1,000	0,962	0,986	1,000
	20	1,000	1,000	0,592	0,857	1,000
	25	1,000	1,000	0,337	0,693	1,000
	50	1,000	1,000	0,096	0,297	1,000
150	8	0,995	1,000	1,000	1,000	1,000
	10	1,000	1,000	1,000	1,000	1,000
	12	1,000	1,000	0,999	1,000	1,000
	15	1,000	1,000	0,994	0,997	1,000
	20	1,000	1,000	0,750	0,949	1,000
	25	1,000	1,000	0,466	0,858	1,000
	50	1,000	1,000	0,100	0,403	1,000

Simülasyon sonucunda (Tablo 3–7) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizine alınan çalışma sayısı  $k=8$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Schwarzer testine yakın bir performansla Harbord testi ve daha sonra Thompson testi takip etmektedir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg ve Egger testleridir. Bu iki testin benzer sonuçlar verdiği ve her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir. Schwarzer ve Harbord testlerinin Tip-I hata oranları birbirine ve belirlenen nominal değere yakındır. Schwarzer testi  $OR \leq 2,34$  olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakta fakat  $OR=4,02$  olduğu durumdan itibaren OR değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Schwarzer testi örneklemdeki birim sayısı  $n \geq 20$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır. Harbord testi  $OR \leq 1,50$  olduğu durumlarda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini koruyamasa da iyi performans göstermekte,  $OR=2,34$  olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakta fakat  $OR=4,02$  olduğu durumdan itibaren OR değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Harbord testi örneklemdeki birim sayısı  $n=50$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.  $OR=2,34$  olduğu durumda Harbord ve Schwarzer testleri performansları birbirine eşittir. Thompson testi  $OR=1,00$  Tip-I hata olasılık düzeyini kısmen koruyabilmektedir.  $OR=1,50$  ve  $2,34$  olduğu ve örneklemdeki birim sayısının artması durumlarında Thompson testinin Tip-I hata oranları artarak nominal değerden sapmaktadır.  $OR=4,02$  olduğu durumdan itibaren OR değerinin artması ile, Thompson testinin Tip-I hata oranlarında örneklem hacminin artışı ile nominal değerden sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır.  $OR \leq 2,34$  olduğu durumlarda Begg testinin örneklemdeki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden artış yönünde sapmanın da arttığı gözlenmektedir.  $OR=4,02$  olduğu durumdan itibaren OR değerinin artması



ile, Begg testinin Tip-I hata oranlarında örneklem hacminin artışı ile nominal değerden artış yönünde sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır.  $OR \leq 4,02$  olduğu durumlarda Egger testinin örneklemdeki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden artış yönünde sapmanın da arttığı gözlenmektedir.  $OR=9,01$  olduğu durumdan itibaren OR değerinin artması ile Egger testinin Tip-I hata oranlarında örneklem hacminin artışı ile nominal değerden artış yönünde sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır.

Simülasyon sonucunda (Tablo 3–7) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizine alınan çalışma sayısı  $k=10$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Schwarzer testine yakın bir performansla Harbord testi takip etmektedir.  $OR=1,00$  ve  $2,34$  olduğu durumlarda Harbord ve Schwarzer testleri performansları birbirine eşit,  $OR=4,02$  olduğu durumda ise Harbord testi Schwarzer testinden daha iyi bir performans göstermiştir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg, Egger ve Thompson testleridir. Bu üç test her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etmekle birlikte, Begg ve Egger testleri Tip-I hata oranları birbirine benzer sonuçlar vermiş ve OR değerinin artması ile Thompson testi Tip-I hata oranlarından daha fazla artış göstererek nominal değerden artış yönünde daha fazla sapsmışlardır.  $OR \leq 2,34$  olduğu durumlarda Begg ve Thompson testlerinin örneklemdeki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden sapmanın da arttığı gözlenmektedir.  $OR=4,02$  olduğu durumdan itibaren OR değerinin artması ile Begg Thompson testlerinin Tip-I hata oranlarında örneklem hacminin artışı ile nominal değerden sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır.  $OR \leq 4,02$  olduğu durumlarda Egger testinin örneklemdeki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden sapmanın da arttığı gözlenmektedir.  $OR=9,01$  olduğu durumdan itibaren OR değerinin artması ile Egger testinin Tip-I hata oranlarında örneklem hacminin artışı ile nominal değerden sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır. Schwarzer ve Harbord testlerinin

Tip-I hata oranları birbirine ve belirlenen nominal değere yakındır. Schwarzer testi  $OR=1,50$  ve  $2,34$  olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakta fakat  $OR=4,02$  olduğu durumdan itibaren  $OR$  değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Schwarzer testi örneklemdaki birim sayısı  $n \geq 25$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır. Harbord testi  $OR \leq 1,50$  ve  $4,02$  olduğu durumlarda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini koruyamasa da iyi performans göstermekte,  $OR=2,34$  olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakta fakat  $OR=9,01$  olduğu durumdan itibaren  $OR$  değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde oldukça sapmalı olarak tahmin etme eğilimindedir. Harbord testi örneklemdaki birim sayısı  $n=50$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.

Simülasyon sonucunda (Tablo 3–7) örneklem hacmi ve  $OR$  değerleri ayrımı yapılmaksızın meta analizine alınan çalışma sayısı  $k=15$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Schwarzer testine yakın bir performansla Harbord testi takip etmektedir.  $OR=1,00$  olduğu durumda Harbord ve Schwarzer testleri performansları birbirine eşit,  $OR=2,34$  olduğu durumda ise Harbord testi Schwarzer testinden daha iyi bir performans göstermiştir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg, Egger ve Thompson testleridir. Bu üç test her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etmekle birlikte, Begg ve Egger testleri Tip-I hata oranları birbirine benzer sonuçlar vermiş ve  $OR$  değerinin artması ile Thompson testi Tip-I hata oranlarından daha fazla artış göstererek nominal değerden daha fazla sapmışlardır.  $OR \leq 2,34$  olduğu durumlarda Begg, Egger ve Thompson testlerinin örneklemdaki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden artış yönünde sapmanın da arttığı gözlenmektedir.  $OR=4,02$  olduğu durumdan itibaren  $OR$  değerinin artması ile Begg Egger ve Thompson testlerinin Tip-I hata oranlarında örneklem hacminin artışı ile nominal değerden sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma

azalmaktadır. Schwarzer ve Harbord testlerinin Tip-I hata oranları birbirine ve belirlenen nominal değere yakındır. Schwarzer testi  $OR=1,50$  olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakta fakat  $OR=2,34$  olduğu durumdan itibaren OR değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Schwarzer testi örneklemdaki birim sayısı  $n \geq 25$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır. Harbord testi  $OR \leq 2,34$  olduğu durumlarda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini koruyamasa da iyi performans göstermekte, fakat  $OR=4,02$  olduğu durumdan itibaren OR değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde oldukça sapmalı olarak tahmin etme eğilimindedir. Harbord testi örneklemdaki birim sayısı  $n=50$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.

Simülasyon sonucunda (Tablo 3–7) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizine alınan çalışma sayısı  $k=20$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Schwarzer testine yakın bir performansla Harbord testi takip etmektedir.  $OR=1,50$  ve  $2,34$  olduğu durumlarda Harbord ve Schwarzer testleri performansları birbirine eşittir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg, Egger ve Thompson testleridir. Bu üç test her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etmiş ve bu testlerin Tip-I hata oranları birbirine benzer sonuçlar vermiştir.  $OR \leq 2,34$  olduğu durumlarda Begg testinin örneklemdaki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden artış yönünde sapmanın da arttığı gözlenmektedir.  $OR=4,02$  olduğu durumdan itibaren OR değerinin artması ile, Begg testinin Tip-I hata oranlarında örneklem hacminin artışı ile nominal değerden sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır.  $OR \leq 4,02$  olduğu durumlarda Egger ve Thompson testlerinin örneklemdaki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden artış yönünde sapmanın da arttığı gözlenmektedir.  $OR=9,01$  olduğu durumdan itibaren OR değerinin artması ile Egger ve Thompson testlerinin Tip-I hata oranlarında örneklem hacminin artışı ile nominal değerden sapma

artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır. Schwarzer ve Harbord testlerinin Tip-I hata oranları birbirine ve belirlenen nominal değere yakındır. Schwarzer testi OR=1,50 ve 2,34 olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakta fakat OR=4,02 olduğu durumdan itibaren OR değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Schwarzer testi örneklemdeki birim sayısı  $n \geq 25$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır. Harbord testi OR=1,50 ve 2,34 olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakta fakat OR=4,02 olduğu durumdan itibaren OR değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde oldukça sapmalı olarak tahmin etme eğilimindedir. Harbord testi örneklemdeki birim sayısı  $n=50$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.

Simülasyon sonucunda (Tablo 3–7) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizine alınan çalışma sayısı  $k=25$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Schwarzer testine yakın bir performansla Harbord testi takip etmektedir. OR=2,34 olduğu durumlarda Harbord ve Schwarzer testleri performansları birbirine eşittir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg, Egger ve Thompson testleridir. Bu üç test her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etmiş ve bu testlerin Tip-I hata oranları birbirine benzer sonuçlar vermiştir. OR $\leq$ 2,34 olduğu durumlarda Begg testinin örneklemdeki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden artış yönünde sapmanın da arttığı gözlenmektedir. Begg testinin Tip-I hata oranlarında OR=4,02 olduğu durumdan itibaren OR değeri arttıkça, örneklem hacminin artışı ile nominal değerden sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır. OR $\leq$ 4,02 olduğu durumlarda Egger ve Thompson testlerinin örneklemdeki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden artış yönünde sapmanın da arttığı gözlenmektedir. OR=9,01 olduğu durumdan itibaren OR değerinin artması ile, Egger ve Thompson testlerinin Tip-I hata oranlarında örneklem

hacminin artışı ile nominal değerden sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır. Schwarzer ve Harbord testlerinin Tip-I hata oranları birbirine ve belirlenen nominal değere yakındır. Schwarzer testi  $OR \leq 2,34$  olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakta fakat  $OR=4,02$  olduğu durumdan itibaren OR değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Schwarzer testi örneklemdeki birim sayısı  $n=50$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır. Harbord testi  $OR=2,34$  olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakta fakat  $OR=4,02$  olduğu durumdan itibaren OR değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde oldukça sapmalı olarak tahmin etme eğilimindedir. Harbord testi  $OR=9,01$  olduğu her durumda Tip-I hata düzeyini sapmalı olarak tahmin etmiştir.

Simülasyon sonucunda (Tablo 3–7) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizine alınan çalışma sayısı  $k=50$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Schwarzer testine yakın bir performansla Harbord testi takip etmektedir.  $OR=1,00$  olduğu durumda Harbord ve Schwarzer testleri performansları birbirine eşit,  $OR=1,50$  olduğu durumda ise Harbord testi Schwarzer testinden daha iyi bir performans göstermiştir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg, Egger ve Thompson testleridir. Bu üç test her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etmiş ve bu testlerin Tip-I hata oranları birbirine benzer sonuçlar vermiştir.  $OR \leq 4,02$  olduğu durumlarda Begg, Egger ve Thompson testlerinin örneklemdeki birim sayısının artması ile Tip-I hata oranlarında nominal düzeyden sapmanın da arttığı gözlenmektedir.  $OR=9,01$  olduğu durumdan itibaren OR değerinin artması ile, Begg, Egger ve Thompson testlerinin Tip-I hata oranlarında örneklem hacminin artışı ile nominal değerden sapma artmakta, örneklem hacminin azalması ile nominal değerden sapma azalmaktadır. Schwarzer ve Harbord testlerinin Tip-I hata oranları birbirine ve belirlenen nominal değere yakındır. Schwarzer testi  $OR=2,34$  olduğu her durumda tam olarak başlangıçta

belirlenen Tip-I hata olasılık deęerini korumakta fakat  $OR=4,02$  olduęu durumdan itibaren  $OR$  deęeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Schwarzer testi örneklemdeki birim sayısı  $n=50$  olduęu her durumda başlangıçta belirlenen Tip-I hata olasılık deęerini korumaktadır. Harbord testi her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık deęerini koruyamamakla birlikte  $OR \leq 2,34$  olduęu durumlarda iyi performans göstermektedir. Harbord testi  $OR=2,34$  olduęu durumdan itibaren  $OR$  deęeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Harbord testi  $OR=9,01$  olduęu her durumda Tip-I hata düzeyini sapmalı olarak tahmin etmiştir.

Simülasyon sonucunda (Tablo 3–7) örneklem hacmi ve  $OR$  deęerleri ayrımı yapılmaksızın meta analizine alınan çalışma sayısı  $k=100$  olduęu durumda genel olarak Tip-I hata oranları nominal deęere göre deęerlendirildięinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık deęerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Harbord testi göstermiştir. Bu testi, Harbord testine yakın bir performansla Schwarzer testi takip etmektedir.  $OR=1,50$  olduęu durumda Harbord testi Schwarzer testinden daha iyi bir performans göstermiştir. Dięer durumlarda Schwarzer testinin performansı Harbord testinden daha iyidir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg, Egger ve Thompson testleridir. Bu üç test her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etmiş ve bu testlerin Tip-I hata oranları birbirine benzer sonuçlar vermiştir.  $OR$  deęerinin artması ile Begg, Egger ve Thompson testlerinin Tip-I hata oranlarında nominal deęerden sapma artmaktadır. Schwarzer ve Harbord testlerinin Tip-I hata oranları birbirine yakındır. Schwarzer testi her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık deęerini koruyamamakla birlikte  $OR=1,00$  ve  $2,34$  olduęu durumlarda iyi performans göstermektedir. Schwarzer testi  $OR=4,02$  olduęu durumdan itibaren  $OR$  deęeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Schwarzer testi örneklemdeki birim sayısı  $n=50$  olduęu her durumda başlangıçta belirlenen Tip-I hata olasılık deęerini korumaktadır. Harbord testi  $OR=1,50$  olduęu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık deęerini korumakla birlikte  $OR=1,00$  ve  $2,34$  olduęu durumlarda iyi performans

göstermektedir. Harbord testi  $OR=2,34$  olduğu durumdan itibaren  $OR$  değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Harbord testi  $OR=9,01$  olduğu her durumda Tip-I hata düzeyini sapmalı olarak tahmin etmiştir.

Simülasyon sonucunda (Tablo 3–7) örneklem hacmi ve  $OR$  değerleri ayrımı yapılmaksızın meta analizine alınan çalışma sayısı  $k=150$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Harbord testi göstermiştir. Bu testi, Harbord testine yakın bir performansla Schwarzer testi takip etmektedir.  $OR=1,00$  ve  $1,50$  olduğu durumlarda Harbord testi Schwarzer testinden daha iyi bir performans göstermiştir. Diğer durumlarda Schwarzer testinin performansı Harbord testinden daha iyidir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg, Egger ve Thompson testleridir. Bu üç test her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etmiş ve bu testlerin Tip-I hata oranları birbirine benzer sonuçlar vermiştir.  $OR$  değerinin artması ile Begg, Egger ve Thompson testlerinin Tip-I hata oranlarında nominal değerden sapma artmaktadır. Schwarzer ve Harbord testlerinin Tip-I hata oranları birbirine yakındır. Schwarzer testi her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini koruyamamakla birlikte  $OR=2,34$  olduğu durumda iyi performans göstermektedir denilebilir. Schwarzer testi  $OR=4,02$  olduğu durumdan itibaren  $OR$  değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Schwarzer testi örneklemdeki birim sayısı  $n=50$  olduğu her durumda başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır. Harbord testi  $OR=1,50$  olduğu her durumda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumakla birlikte  $OR=1,00$  ve  $2,34$  olduğu durumlarda iyi performans göstermektedir. Harbord testi  $OR=2,34$  olduğu durumdan itibaren  $OR$  değeri arttıkça ve örneklem hacmi azaldıkça Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğilimindedir. Harbord testi  $OR=9,01$  olduğu her durumda Tip-I hata düzeyini sapmalı olarak tahmin etmiştir.

Simülasyon sonucunda (Tablo 3–7) meta analizine alınan çalışma sayısı ve  $OR$  değerleri ayrımı yapılmaksızın örneklem hacmi  $n=8$  olduğu durumda genel olarak Tip-

I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Harbord testi ve daha sonra Thompson testi takip etmektedir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg ve Egger testleridir. Bu iki testin benzer sonuçlar verdiği ve her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir.  $OR \leq 2,34$  olduğu durumlarda Schwarzer testi, Harbord testinden daha iyi performans göstermiştir. Harbord ve Schwarzer testlerinin  $OR \geq 4,02$  olduğu her durumda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir. Harbord testi  $OR \leq 1,50$  olduğu durumlarda Tip-I hata düzeyini azalış yönünde sapmalı olarak tahmin etme eğilimindedir.

Simülasyon sonucunda (Tablo 3–7) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın örneklem hacmi  $n=10$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Harbord testi ve daha sonra Thompson testi takip etmektedir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg ve Egger testleridir. Bu iki testin benzer sonuçlar verdiği ve her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir.  $OR=1,00$  ve  $2,34$  olduğu durumlarda Schwarzer testi, Harbord testinden daha iyi performans göstermiştir.  $OR=1,50$  olduğu durumda Harbord ve Schwarzer testleri aynı performansı göstermiştir.  $OR=4,02$  olduğu durumda Harbord testi, Schwarzer testinden daha iyi performans göstermiştir. Harbord ve Schwarzer testleri  $OR=9,01$  olduğu her durumda Tip-I hata düzeyini artış önünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir. Harbord testi  $OR \leq 1,50$  olduğu durumlarda Tip-I hata düzeyini azalış yönünde sapmalı olarak tahmin etme eğilimindedir.

Simülasyon sonucunda (Tablo 3–7) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın örneklem hacmi  $n=12$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık



değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Harbord testi göstermiştir. Bu testi, Harbord testine yakın bir performansla Schwarzer testi ve daha sonra Thompson testi takip etmektedir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg ve Egger testleridir. Bu iki testin benzer sonuçlar verdiği ve her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir.  $OR=4,02$ , olduğu durumlarda Schwarzer testi Harbord testinden daha iyi performans göstermiştir.  $OR=1,50$  ve  $2,34$  olduğu durumlarda Harbord ve Schwarzer testleri aynı performansı göstermiştir.  $OR=1,00$  olduğu durumda Harbord testi Schwarzer testinden daha iyi performans göstermiştir. Harbord ve Schwarzer testleri  $OR=9,01$  olduğu her durumda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir.

Simülasyon sonucunda (Tablo 3–7) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın örneklem hacmi  $n=15$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Harbord testi göstermiştir. Bu testi, Harbord testine yakın bir performansla Schwarzer testi takip etmektedir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Thompson, Begg ve Egger testleridir. Bu testlerin her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir.  $OR=4,02$  olduğu durumda Schwarzer testi Harbord testinden daha iyi performans göstermiştir.  $OR=2,34$  olduğu durumda Harbord ve Schwarzer testleri aynı performansı göstermiştir.  $OR \leq 1,50$  olduğu durumlarda Harbord testi Schwarzer testinden daha iyi performans göstermiştir. Harbord ve Schwarzer testlerinin de diğer üç test gibi  $OR=9,01$  olduğu her durumda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir.  $OR \leq 1,50$  olduğu her durumda Harbord testi tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.

Simülasyon sonucunda (Tablo 3–7) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın örneklem hacmi  $n=20$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Harbord testi

göstermiştir. Bu testi, Harbord testine yakın bir performansla Schwarzer testi takip etmektedir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Thompson, Begg ve Egger testleridir. Bu testlerin her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir.  $OR \geq 4,02$  olduğu durumlarda Schwarzer testi Harbord testinden daha iyi performans göstermiştir.  $OR=2,34$  olduğu durumda Harbord ve Schwarzer testleri aynı performansı göstermiştir.  $OR \leq 1,50$  olduğu durumlarda Harbord testi Schwarzer testinden daha iyi performans göstermiştir. Harbord testi  $OR=9,01$  olduğu her durumda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir.  $OR=2,34$  ve  $4,02$  olduğu her durumda Schwarzer testi tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.  $OR \leq 2,34$  olduğu her durumda Harbord testi tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.

Simülasyon sonucunda (Tablo 3–7) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın örneklem hacmi  $n=25$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Harbord testi takip etmektedir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Thompson, Begg ve Egger testleridir. Bu testlerin her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir.  $OR \geq 4,02$  olduğu durumlarda Schwarzer testi Harbord testinden daha iyi performans göstermiştir.  $OR=1,00$  olduğu durumda Harbord ve Schwarzer testleri aynı performansı göstermiştir.  $OR=1,50$  ve  $2,34$  olduğu durumda Harbord testi Schwarzer testinden daha iyi performans göstermiştir.  $OR=1,00$  ve  $4,02$  olduğu her durumda Schwarzer testi tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.  $OR \leq 2,34$  olduğu her durumda Harbord testi tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.

Simülasyon sonucunda (Tablo 3–7) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın örneklem hacmi  $n=50$  olduğu durumda genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında

kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı Schwarzer testi göstermiştir. Bu testi, Schwarzer testine yakın bir performansla Harbord testi ve daha sonra Thompson testi takip etmektedir. Tip-I hatayı tahmin etme yönünden en kötü performansı gösteren testler Begg ve Egger testleridir. Bu testlerin her koşulda Tip-I hata düzeyini artış yönünde sapmalı olarak tahmin etme eğiliminde olduğu görülmektedir. OR=9,01 olduğu durumlarda Schwarzer testi, Harbord testinden daha iyi performans göstermiştir. OR=1,50 olduğu durumda Harbord testi Schwarzer testinden daha iyi performans göstermiştir. OR=1,50 olduğu durum hariç diğer her durumda Schwarzer testi tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır. OR $\leq$ 4,02 olduğu her durumda Harbord testi tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini korumaktadır.

Meta analizinde yanlılığın olmadığı durum dikkate alınarak yapılan simülasyon senaryoları (Tablo 3–7) incelendiğinde örneklem hacmi, meta analize alınan çalışma sayıları ve OR değeri ayırımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, bu tez çalışmasında kullanılan testlerin her koşulda tam olarak başlangıçta belirlenen Tip-I hata olasılık değerini ( $\alpha=0,10$ ) koruyamamasıyla birlikte, en iyi performansı gösteren test Schwarzer testidir. Schwarzer testi dışında Harbord testi de ele alınan kombinasyonlar bakımından Schwarzer testine yakın bir performansla Tip-I hata düzeyini koruma eğilimindedir. Thompson testi ise sadece  $k=8$  gibi meta analize alınan çalışma sayısı az olduğu birkaç durumda Tip-I hatayı koruma yönünde performans göstermiştir. Begg ve Egger testlerinin performanslarının genel olarak kötü olduğu ve Tip-I hatayı koruma yönünde başarısız oldukları söylenebilir. Meta analize alınan çalışma sayılarının artması ile bu tez çalışmasında kullanılan bütün testlerin Tip-I hata oranları artmıştır. OR değerinin artmasıyla Begg, Egger, Thompson testlerinde Tip-I hata oranlarının ve dolayısıyla nominal değerden sapmanın önce arttığı, sonra azaldığı saptanmıştır. OR değerinin artmasıyla Schwarzer ve Harbord testlerinde örneklem hacminin azalması ve meta analize alınan çalışma sayılarının artması ile Tip-I hata oranlarının ve dolayısıyla nominal değerden sapmanın oldukça arttığı görülmektedir. Genel olarak örneklem hacminin artması ile Schwarzer ve Harbord testlerinin Tip-I hatayı koruma kriterlerine göre gösterdikleri performans yükselmiştir. Genel olarak

örneklem hacminin artması ile Begg, Egger, Thompson testlerinde Tip-I hata oranlarının ve dolayısıyla nominal değerden artış yönünde sapmanın artma eğiliminde olduğu söylenebilir. OR değeri 1 değerine yakın olduğunda meta analizine alınan çalışma sayısı ve örneklem hacmi az olduğunda Begg, Egger, Thompson testleri her ne kadar iyi performans göstermeseler de Tip-I hata oranları, nominal değere daha yakındır. Ek olarak Begg ve Egger testlerinin Tip-I hatayı koruma kriterlerine göre gösterdikleri performanslar incelendiğinde genel olarak liberal bir tutum gösterdikleri, Tip-I hata oranını nominal düzeyinin üstünde tahmin etme eğiliminde oldukları söylenebilir.

#### **4.2. Meta Analizde %10 Yanlılık Oluşturduğu Duruma Ait Bulgular**

Meta analizine alınan çalışmaların OR değerleri sıralanıp en büyük OR değerlerine sahip olanların %10'u meta analizinden çıkarılarak funnel grafiğinin sağ tarafında asimetri sağlanması ile yanlılık oluşturulduğu durum dikkate alınarak yapılan simülasyon çalışması ile elde edilen veriler analiz edilmiş ve analizlerden elde edilen bulgular tablolar (Tablo 8 – 12) aracılığıyla sunulmuştur. Hasta grubunda pozitif sonuç oranlarına göre ayrı ayrı sunulan tablolarda anakütle OR değerleri,  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  olan örneklem hacimleri ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  olan meta analizine alınan çalışma sayılarına göre funnel grafiği asimetrisinin değerlendirilmesi ile meta analizinde yanlılık olup olmadığına karar vermemizi sağlayan Begg, Egger, Schwarzer, Harbord ve Thompson testlerine ait yanlılığı tespit etme güçleri verilmiştir. Nominal değer yani Tip-I hata düzeyi  $\alpha=0,10$  olarak alınmıştır.

Meta analizinde %10 yanlılık oluşturulduğu,  $OR=1,00$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,50$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 8'de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi ve Thompson testi göstermiştir. Egger ve Thompson testlerini, yakın bir performansla Begg testi takip etmektedir. Begg, Egger

ve Thompson testlerinin yanlılıđı tespit etme gcnn meta analizine alınan alıřma sayısı arttıķa ykseldiđi zellikle  $k \geq 100$  olduđu durumlarda %80'nin zerine ıktıđı grlmektedir. Az sayıda alıřma ieren meta analizler iin bu tez alıřmasında kullanılan btn testlerin gc olduka dřktr. Harbord ve Schwarzer testleri en az gce sahiptir ve alıřma sayısının artıřından etkilenmemiřtir. Begg testinin yanlılıđı tespit etme gc bakımından performansı rneklem hacminin artmasıyla dřme eđilimindedir. Begg testi hari bu tez alıřmasında incelenen diđer testler genel olarak rneklem hacminin deđiřiminden etkilenmemiřtir.

**Tablo 8.** Meta analizinde %10 yanlışlık oluşturulduğu durumda OR=1,00  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,50$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

<b>k</b>	<b><math>n_H=n_K</math></b>	<b>Begg Yöntemi</b>	<b>Egger Yöntemi</b>	<b>Schwarzer Yöntemi</b>	<b>Harbord Yöntemi</b>	<b>Thompson Yöntemi</b>
<b>8</b>	<b>8</b>	0,278	0,316	0,100	0,087	0,081
	<b>10</b>	0,283	0,344	0,122	0,086	0,110
	<b>12</b>	0,280	0,335	0,129	0,102	0,093
	<b>15</b>	0,306	0,360	0,117	0,117	0,118
	<b>20</b>	0,256	0,301	0,106	0,091	0,093
	<b>25</b>	0,269	0,314	0,106	0,093	0,093
<b>10</b>	<b>50</b>	0,227	0,334	0,102	0,107	0,102
	<b>8</b>	0,294	0,361	0,132	0,077	0,134
	<b>10</b>	0,267	0,364	0,121	0,081	0,149
	<b>12</b>	0,282	0,341	0,131	0,103	0,156
	<b>15</b>	0,317	0,367	0,110	0,100	0,175
	<b>20</b>	0,244	0,310	0,098	0,086	0,141
<b>15</b>	<b>25</b>	0,259	0,353	0,099	0,094	0,153
	<b>50</b>	0,232	0,310	0,081	0,090	0,132
	<b>8</b>	0,386	0,475	0,118	0,067	0,276
	<b>10</b>	0,377	0,485	0,119	0,087	0,310
	<b>12</b>	0,376	0,462	0,130	0,094	0,265
	<b>15</b>	0,384	0,483	0,122	0,090	0,290
<b>20</b>	<b>20</b>	0,374	0,459	0,109	0,091	0,279
	<b>25</b>	0,380	0,459	0,126	0,101	0,279
	<b>50</b>	0,320	0,411	0,086	0,102	0,252
	<b>8</b>	0,374	0,477	0,147	0,082	0,341
	<b>10</b>	0,356	0,486	0,124	0,075	0,342
	<b>12</b>	0,388	0,483	0,120	0,096	0,361
<b>25</b>	<b>15</b>	0,384	0,507	0,122	0,096	0,368
	<b>20</b>	0,386	0,484	0,104	0,085	0,337
	<b>25</b>	0,342	0,461	0,101	0,090	0,324
	<b>50</b>	0,321	0,455	0,098	0,102	0,313
	<b>8</b>	0,463	0,538	0,159	0,075	0,400
	<b>10</b>	0,469	0,602	0,129	0,075	0,484
<b>50</b>	<b>12</b>	0,464	0,568	0,133	0,086	0,434
	<b>15</b>	0,478	0,603	0,115	0,100	0,457
	<b>20</b>	0,464	0,608	0,115	0,095	0,449
	<b>25</b>	0,446	0,573	0,101	0,112	0,427
	<b>50</b>	0,391	0,540	0,111	0,103	0,406
	<b>8</b>	0,588	0,676	0,153	0,071	0,577
<b>100</b>	<b>10</b>	0,599	0,763	0,134	0,064	0,678
	<b>12</b>	0,566	0,741	0,129	0,087	0,657
	<b>15</b>	0,625	0,761	0,132	0,093	0,671
	<b>20</b>	0,586	0,734	0,129	0,100	0,641
	<b>25</b>	0,536	0,717	0,114	0,091	0,617
	<b>50</b>	0,534	0,710	0,116	0,103	0,626
<b>150</b>	<b>8</b>	0,808	0,887	0,144	0,071	0,828
	<b>10</b>	0,805	0,931	0,156	0,077	0,898
	<b>12</b>	0,779	0,924	0,146	0,084	0,893
	<b>15</b>	0,847	0,925	0,151	0,115	0,893
	<b>20</b>	0,786	0,910	0,161	0,106	0,866
	<b>25</b>	0,751	0,911	0,143	0,097	0,869
<b>150</b>	<b>50</b>	0,730	0,900	0,123	0,090	0,854
	<b>8</b>	0,919	0,958	0,166	0,067	0,927
	<b>10</b>	0,922	0,986	0,201	0,078	0,976
	<b>12</b>	0,890	0,969	0,180	0,104	0,954
	<b>15</b>	0,922	0,976	0,150	0,105	0,963
	<b>20</b>	0,896	0,972	0,165	0,104	0,957
<b>150</b>	<b>25</b>	0,879	0,970	0,151	0,088	0,954
	<b>50</b>	0,850	0,971	0,121	0,104	0,951

Meta analizinde %10 yanlılık oluşturulduğu,  $OR=1,50$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,60$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 9’da gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Begg testi göstermiştir. Begg testini, bu teste çok yakın bir performansla Egger testi ve daha sonra Thompson testi takip etmektedir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücünün meta analizine alınan çalışma sayısı arttıkça yükseldiği, Begg ve Egger testlerinin özellikle  $k \geq 15$ , Thompson testinin ise  $k \geq 20$  olduğu durumlarda güçlerinin %80’nin üzerine çıktığı görülmektedir. Schwarzer testinin yanlılık belirleme gücü meta analizine alınan çalışma sayısı fazla ise (50, 100, 150 gibi) az da olsa yükselmeye başlamıştır. Harbord testi en az güce sahiptir ve meta analizine alınan çalışma sayısından ve örneklem hacminin değişiminden etkilenmemiştir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü örnekleme alınan birim sayısı arttıkça yükselmiştir. Harbord ve Schwarzer testlerinin gücü örneklem hacminin değişiminden etkilenmemiştir. Az sayıda çalışma içeren meta analizler için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür. OR değerinin artışıyla birlikte Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları oldukça yükselmiştir. OR değerinin artışıyla birlikte Schwarzer testinin gücünde az da olsa bir yükseliş meydana gelmiş, Harbord testinin gücünde belirgin bir değişim olmamıştır.

**Tablo 9.** Meta analizinde %10 yanlışlık oluşturulduğu durumda  $OR=1,50$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,60$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,526	0,528	0,124	0,083	0,168
	10	0,556	0,573	0,122	0,076	0,188
	12	0,586	0,600	0,114	0,091	0,188
	15	0,617	0,587	0,089	0,084	0,207
	20	0,629	0,614	0,127	0,107	0,223
	25	0,664	0,605	0,112	0,099	0,208
	50	0,654	0,663	0,101	0,106	0,264
10	8	0,561	0,558	0,118	0,074	0,249
	10	0,623	0,628	0,121	0,077	0,323
	12	0,641	0,640	0,113	0,077	0,311
	15	0,710	0,655	0,102	0,095	0,342
	20	0,722	0,676	0,112	0,099	0,358
	25	0,740	0,679	0,110	0,096	0,366
	50	0,752	0,735	0,100	0,094	0,425
15	8	0,797	0,759	0,091	0,066	0,507
	10	0,851	0,791	0,134	0,078	0,587
	12	0,863	0,841	0,103	0,069	0,620
	15	0,908	0,833	0,118	0,080	0,602
	20	0,917	0,841	0,134	0,097	0,611
	25	0,938	0,871	0,122	0,097	0,634
	50	0,935	0,876	0,123	0,108	0,668
20	8	0,812	0,774	0,118	0,071	0,618
	10	0,901	0,843	0,134	0,071	0,721
	12	0,908	0,878	0,117	0,082	0,761
	15	0,960	0,907	0,142	0,100	0,772
	20	0,968	0,906	0,128	0,088	0,785
	25	0,975	0,912	0,123	0,101	0,797
	50	0,981	0,946	0,112	0,106	0,837
25	8	0,932	0,871	0,115	0,059	0,762
	10	0,968	0,934	0,135	0,065	0,838
	12	0,963	0,945	0,128	0,086	0,866
	15	0,989	0,956	0,152	0,097	0,879
	20	0,989	0,954	0,132	0,082	0,861
	25	0,993	0,958	0,153	0,094	0,869
	50	0,993	0,969	0,127	0,104	0,903
50	8	0,999	0,988	0,121	0,058	0,967
	10	1,000	0,990	0,163	0,056	0,984
	12	0,999	0,996	0,187	0,076	0,985
	15	1,000	0,997	0,213	0,110	0,992
	20	1,000	0,996	0,184	0,082	0,992
	25	1,000	0,999	0,198	0,121	0,998
	50	1,000	0,998	0,127	0,098	0,995
100	8	1,000	1,000	0,158	0,072	1,000
	10	1,000	1,000	0,282	0,059	1,000
	12	1,000	1,000	0,292	0,083	1,000
	15	1,000	1,000	0,296	0,113	1,000
	20	1,000	1,000	0,287	0,101	1,000
	25	1,000	1,000	0,257	0,107	1,000
	50	1,000	1,000	0,178	0,090	1,000
150	8	1,000	1,000	0,188	0,078	1,000
	10	1,000	1,000	0,364	0,061	1,000
	12	1,000	1,000	0,424	0,072	1,000
	15	1,000	1,000	0,378	0,109	1,000
	20	1,000	1,000	0,399	0,101	1,000
	25	1,000	1,000	0,347	0,130	1,000
	50	1,000	1,000	0,228	0,094	1,000



Meta analizinde %10 yanlılık oluşturulduğu,  $OR=2,34$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,70$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 10'da gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayırımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Begg ve Egger testleri göstermiştir. Begg ve Egger testlerini, bu testlere yakın bir performansla Thompson testi takip etmektedir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücünün meta analizine alınan çalışma sayısı arttıkça yükseldiği, Begg ve Egger testlerinin özellikle  $k \geq 15$ , Thompson testinin ise  $k \geq 20$  olduğu her durumda güçlerinin %80'nin üzerine çıktığı görülmektedir. Schwarzer ve Harbord testleri en az güce sahiptir ve çalışma sayısının ve örneklem hacminin değişiminden etkilenmemiştir. Az sayıda çalışma içeren meta analizler için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı örnekleme alınan birim sayısı arttıkça yükselme eğiliminde iken, Harbord ve Schwarzer testlerinin gücü örneklem hacminin değişiminden etkilenmemiştir. OR değerinin artışıyla birlikte Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları artmıştır. OR değerinin artışıyla birlikte Schwarzer testinin gücünde düşüş meydana gelmiş, Harbord testinin gücünde belirgin bir değişim olmamıştır.

**Tablo 10.** Meta analizinde %10 yanlışlık oluşturulduğu durumda  $OR=2,34$   $P(S^+|H^+)=0,50$  ve  $P(S^+|H^-)=0,70$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

<b>k</b>	<b><math>n_H=n_K</math></b>	<b>Begg Yöntemi</b>	<b>Egger Yöntemi</b>	<b>Schwarzer Yöntemi</b>	<b>Harbord Yöntemi</b>	<b>Thompson Yöntemi</b>
<b>8</b>	<b>8</b>	0,634	0,648	0,103	0,068	0,187
	<b>10</b>	0,689	0,691	0,088	0,074	0,243
	<b>12</b>	0,678	0,701	0,093	0,080	0,232
	<b>15</b>	0,680	0,698	0,089	0,103	0,244
	<b>20</b>	0,644	0,702	0,112	0,096	0,254
	<b>25</b>	0,658	0,694	0,112	0,119	0,274
<b>10</b>	<b>8</b>	0,708	0,694	0,092	0,068	0,337
	<b>10</b>	0,800	0,780	0,093	0,064	0,420
	<b>12</b>	0,775	0,765	0,096	0,072	0,404
	<b>15</b>	0,793	0,797	0,111	0,093	0,437
	<b>20</b>	0,777	0,805	0,113	0,104	0,438
	<b>25</b>	0,769	0,787	0,094	0,092	0,453
<b>15</b>	<b>8</b>	0,917	0,865	0,079	0,053	0,580
	<b>10</b>	0,956	0,925	0,078	0,054	0,663
	<b>12</b>	0,935	0,917	0,090	0,070	0,700
	<b>15</b>	0,944	0,933	0,104	0,081	0,712
	<b>20</b>	0,947	0,929	0,111	0,096	0,726
	<b>25</b>	0,941	0,938	0,090	0,092	0,721
<b>20</b>	<b>8</b>	0,971	0,923	0,091	0,077	0,758
	<b>10</b>	0,978	0,962	0,077	0,053	0,849
	<b>12</b>	0,985	0,967	0,109	0,073	0,862
	<b>15</b>	0,984	0,970	0,104	0,087	0,878
	<b>20</b>	0,980	0,976	0,119	0,098	0,884
	<b>25</b>	0,984	0,979	0,098	0,094	0,886
<b>25</b>	<b>8</b>	0,974	0,971	0,110	0,106	0,884
	<b>10</b>	0,990	0,970	0,081	0,070	0,854
	<b>12</b>	0,995	0,989	0,088	0,062	0,934
	<b>15</b>	0,999	0,989	0,081	0,050	0,940
	<b>20</b>	0,997	0,992	0,107	0,077	0,948
	<b>25</b>	0,995	0,988	0,126	0,103	0,934
<b>50</b>	<b>8</b>	0,998	0,988	0,115	0,113	0,939
	<b>10</b>	0,992	0,992	0,101	0,104	0,953
	<b>12</b>	1,000	1,000	0,077	0,104	0,993
	<b>15</b>	1,000	1,000	0,099	0,062	1,000
	<b>20</b>	1,000	1,000	0,116	0,050	1,000
	<b>25</b>	1,000	1,000	0,138	0,075	0,999
<b>100</b>	<b>8</b>	1,000	1,000	0,108	0,067	0,999
	<b>10</b>	1,000	1,000	0,106	0,085	1,000
	<b>12</b>	1,000	1,000	0,102	0,090	0,998
	<b>15</b>	1,000	1,000	0,087	0,169	1,000
	<b>20</b>	1,000	1,000	0,118	0,076	1,000
	<b>25</b>	1,000	1,000	0,154	0,059	1,000
<b>150</b>	<b>8</b>	1,000	1,000	0,178	0,074	1,000
	<b>10</b>	1,000	1,000	0,154	0,093	1,000
	<b>12</b>	1,000	1,000	0,134	0,101	1,000
	<b>15</b>	1,000	1,000	0,095	0,095	1,000
	<b>20</b>	1,000	1,000	0,076	0,267	1,000
	<b>25</b>	1,000	1,000	0,147	0,096	1,000
<b>500</b>	<b>8</b>	1,000	1,000	0,180	0,048	1,000
	<b>10</b>	1,000	1,000	0,234	0,049	1,000
	<b>12</b>	1,000	1,000	0,172	0,086	1,000
	<b>15</b>	1,000	1,000	0,144	0,087	1,000
	<b>20</b>	1,000	1,000	0,094	0,074	1,000
	<b>25</b>	1,000	1,000	0,094	0,074	1,000

Meta analizinde %10 yanlılık oluşturulduğu,  $OR=4,02$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,80$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 11’de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, bu teste çok yakın bir performansla Begg ve Thompson testleri, daha sonra Harbord ve Schwarzer testleri takip etmektedir. Begg, Egger, Thompson, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücünün meta analizine alınan çalışma sayısı arttıkça yükseldiği tespit edilmiştir. Begg, Egger ve Thompson testlerinin güçleri,  $k \leq 25$  iken örneklemdeki birim sayısı arttıkça önce yükselmiş sonra düşmüştür. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı örnekleme alınan birim sayısı arttıkça yükselme eğiliminde iken, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı örnekleme alınan birim sayısı azaldıkça yükselmiştir. OR değerinin artışıyla birlikte Egger, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Begg ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı düşmüştür.

**Tablo 11.** Meta analizinde %10 yanlışlık oluşturulduğu durumda OR=4,02  $P(S^+|H^+)=0,50$  ve  $P(S^+|H^+)=0,80$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,554	0,616	0,152	0,118	0,130
	10	0,614	0,700	0,103	0,097	0,169
	12	0,587	0,682	0,096	0,077	0,160
	15	0,617	0,733	0,096	0,076	0,224
	20	0,673	0,792	0,090	0,070	0,281
	50	0,661	0,777	0,092	0,093	0,320
10	8	0,626	0,669	0,186	0,148	0,234
	10	0,717	0,784	0,112	0,103	0,320
	12	0,716	0,796	0,115	0,081	0,312
	15	0,739	0,841	0,097	0,074	0,410
	20	0,795	0,878	0,088	0,063	0,481
	50	0,783	0,881	0,096	0,090	0,510
15	8	0,791	0,845	0,242	0,228	0,456
	10	0,874	0,926	0,121	0,127	0,538
	12	0,899	0,946	0,085	0,088	0,621
	15	0,933	0,968	0,096	0,061	0,713
	20	0,940	0,974	0,108	0,081	0,769
	50	0,917	0,961	0,120	0,119	0,770
20	8	0,871	0,919	0,291	0,303	0,658
	10	0,959	0,974	0,160	0,175	0,796
	12	0,953	0,979	0,124	0,102	0,840
	15	0,978	0,990	0,095	0,076	0,888
	20	0,988	0,993	0,109	0,079	0,925
	50	0,990	0,998	0,101	0,079	0,935
25	8	0,977	0,989	0,105	0,098	0,924
	10	0,944	0,975	0,335	0,371	0,781
	12	0,980	0,995	0,183	0,220	0,892
	15	0,992	0,996	0,124	0,140	0,931
	20	0,996	0,997	0,104	0,088	0,961
	50	0,998	1,000	0,102	0,075	0,977
50	8	0,995	1,000	0,088	0,071	0,973
	10	0,991	0,999	0,105	0,093	0,972
	12	0,999	0,999	0,519	0,655	0,990
	15	1,000	1,000	0,288	0,412	0,998
	20	1,000	1,000	0,166	0,254	1,000
	50	1,000	1,000	0,117	0,133	1,000
100	8	1,000	1,000	0,110	0,079	1,000
	10	1,000	1,000	0,114	0,083	1,000
	12	1,000	1,000	0,106	0,085	1,000
	15	1,000	1,000	0,799	0,925	1,000
	20	1,000	1,000	0,458	0,681	1,000
	50	1,000	1,000	0,228	0,445	1,000
150	8	1,000	1,000	0,137	0,228	1,000
	10	1,000	1,000	0,113	0,101	1,000
	12	1,000	1,000	0,111	0,084	1,000
	15	1,000	1,000	0,091	0,083	1,000
	20	1,000	1,000	0,930	0,994	1,000
	50	1,000	1,000	0,554	0,828	1,000
150	8	1,000	1,000	0,303	0,608	1,000
	10	1,000	1,000	0,159	0,335	1,000
	12	1,000	1,000	0,113	0,122	1,000
	15	1,000	1,000	0,123	0,097	1,000
	20	1,000	1,000	0,114	0,085	1,000
	50	1,000	1,000	0,114	0,085	1,000

Meta analizinde %10 yanlılık oluşturulduğu,  $OR=9,01$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,90$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 12’de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, Begg, Thompson ve daha sonra Harbord ve Schwarzer testleri takip etmektedir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansının meta analizine alınan çalışma sayısı arttıkça ve örnekleme alınan birim sayısı arttıkça, yükseldiği görülmektedir. Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları meta analizine alınan çalışma sayısı arttıkça ve örnekleme alınan birim sayısı azaldıkça yükselmiştir. OR değerinin artışıyla birlikte Egger, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü yükselmiştir. Begg ve Thompson testlerinde OR değerinin artışıyla birlikte yanlılığı tespit etme gücü azalmıştır.

**Tablo 12.** Meta analizinde %10 yanlışlık oluşturulduğu durumda  $OR=9,01$ ,  $P(S^+|H^+)=0,50$  ve  $P(S^+|H^+)=0,90$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,378	0,506	0,505	0,444	0,036
	10	0,461	0,577	0,353	0,344	0,055
	12	0,455	0,579	0,257	0,281	0,060
	15	0,499	0,689	0,156	0,186	0,096
	20	0,630	0,781	0,109	0,153	0,148
	50	0,700	0,850	0,091	0,115	0,193
10	8	0,435	0,567	0,599	0,546	0,105
	10	0,556	0,676	0,399	0,431	0,134
	12	0,549	0,700	0,307	0,352	0,157
	15	0,639	0,825	0,179	0,221	0,201
	20	0,758	0,902	0,090	0,147	0,317
	50	0,837	0,931	0,093	0,110	0,425
15	8	0,524	0,732	0,759	0,734	0,216
	10	0,690	0,835	0,560	0,645	0,269
	12	0,681	0,837	0,443	0,532	0,296
	15	0,769	0,941	0,242	0,332	0,409
	20	0,927	0,979	0,134	0,215	0,603
	50	0,958	0,989	0,098	0,151	0,734
20	8	0,617	0,809	0,861	0,849	0,384
	10	0,752	0,900	0,650	0,741	0,472
	12	0,801	0,935	0,532	0,626	0,540
	15	0,878	0,986	0,295	0,407	0,685
	20	0,975	0,992	0,162	0,269	0,864
	50	0,992	0,997	0,096	0,181	0,932
25	8	0,690	0,894	0,928	0,938	0,450
	10	0,825	0,948	0,749	0,845	0,601
	12	0,865	0,965	0,619	0,754	0,651
	15	0,930	0,994	0,367	0,514	0,804
	20	0,995	0,999	0,159	0,368	0,936
	50	1,000	1,000	0,105	0,229	0,978
50	8	0,999	0,999	0,104	0,136	0,994
	8	0,894	0,995	0,996	0,997	0,899
	10	0,972	0,999	0,951	0,987	0,965
	12	0,985	1,000	0,872	0,964	0,977
	15	0,998	1,000	0,626	0,802	0,995
	20	1,000	1,000	0,270	0,577	1,000
100	8	0,985	1,000	1,000	1,000	0,997
	10	0,998	1,000	0,999	1,000	1,000
	12	1,000	1,000	0,994	0,999	1,000
	15	1,000	1,000	0,913	0,983	1,000
	20	1,000	1,000	0,454	0,888	1,000
	50	1,000	1,000	0,215	0,695	1,000
150	8	1,000	1,000	0,125	0,279	1,000
	8	1,000	1,000	1,000	1,000	1,000
	10	1,000	1,000	1,000	1,000	1,000
	12	1,000	1,000	0,999	1,000	1,000
	15	1,000	1,000	0,983	0,999	1,000
	20	1,000	1,000	0,628	0,960	1,000
150	25	1,000	1,000	0,245	0,848	1,000
	50	1,000	1,000	0,117	0,361	1,000

Simülasyon sonucunda (Tablo 8-12) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=8$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Begg, Schwarzer, Harbord ve Thompson testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger testi ise sadece  $OR=9,01$  ve  $n \geq 25$  olduğu durumlarda %80 güce ulaşabilmiştir.

Simülasyon sonucunda (Tablo 8-12) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=10$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi takip etmektedir. Schwarzer, Harbord ve Thompson testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger testi  $OR \geq 2,34$  ve  $n \geq 15$  olduğu durumlarda %80 güce ulaşabilmiştir. Begg testi ise  $OR=2,34$  ve  $9,01$  olduğu durumlarda %80 güce ulaşabilmiştir. Meta analizine alınan çalışma sayısı  $k=8$ 'den  $k=10$ 'a çıktığında Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları artmıştır.

Simülasyon sonucunda (Tablo 8-12) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=15$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra çok da iyi olmayan bir performansla Thompson testi takip etmektedir. Schwarzer ve Harbord testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger ve Begg testi  $OR=1,00$  olduğu durum hariç %80 güce ulaşmış, Thompson testi ise sadece  $OR=9,01$  ve  $n=50$  olduğu durumda %80 güce ulaşabilmiştir. Begg testinin yanlılığı tespit etme gücü performansı  $OR=2,34$  değerine kadar yükselmiş, bu OR değerinden sonra özellikle küçük örneklem hacimlerinde

düşmeye başlamıştır. OR=1,50 iken Begg testi Egger testinden iyi performans göstermiş, OR=2,34 iken Egger testi ile aynı performansı göstermiş,  $OR \geq 4,02$  olduğu durumlarda Egger testinden daha kötü performans göstermiştir. Egger testinin yanlılığı tespit etme gücü performansı OR=9,01 değerine kadar yükselmiş, bu OR değerinden sonra özellikle küçük örneklem hacimlerinde düşmeye başlamıştır. Meta analizine alınan çalışma sayısı  $k=10$ 'dan  $k=15$ 'e çıktığında Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları artmıştır.

Simülasyon sonucunda (Tablo 8-12) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=20$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson testi, Schwarzer ve Harbord testleri takip etmektedir. Schwarzer ve Harbord testleri sadece OR=9,01 ve  $n=8$  olduğu durumda %80 güce ulaşabilmiştir. Bu tez çalışmasında incelenen hiçbir test OR=1,00 olduğu durumda %80 güce ulaşamamıştır. Thompson testinin yanlılığı tespit etme gücü bakımından performansı OR=2,34 değerine kadar yükselmiş, daha sonra OR değerinin artışı ile birlikte küçük örneklem hacimlerinde düşmeye başlamıştır. Begg testinin yanlılığı tespit etme gücü OR=1,50, OR=2,34 ve OR=4,02 olduğu her durumda %80'e ulaşmış, OR=9,01 olduğu durumda performansı özellikle küçük örneklem hacimlerinde düşmüştür. Egger testinin yanlılığı tespit etme gücü bakımından performansı OR değeri artıka önce yükselmiş, daha sonra sonra özellikle küçük örneklem hacimlerinde düşmeye başlamıştır. Meta analizine alınan çalışma sayısı  $k=15$ 'ten  $k=20$ 'ye çıktığında bu tez çalışmasında incelenen bütün testlerinin yanlılığı tespit etme gücü bakımından performansları artmıştır.

Simülasyon sonucunda (Tablo 8-12) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=25$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri sadece OR=9,01



ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Bu tez çalışmasında incelenen hiçbir test  $OR=1,00$  olduğu durumda %80 güce ulaşamamıştır. Thompson testinin yanlılığı tespit etme gücü bakımından performansı  $OR=2,34$  değerine kadar yükselmiş, daha sonra  $OR$  değerinin artışı ile birlikte performansı küçük örneklem hacimlerinde düşmeye başlamıştır. Begg ve Egger testlerinin yanlılığı tespit etme gücü  $OR=1,50$ ,  $OR=2,34$  ve  $OR=4,02$  olduğu her durumda %80'e ulaşmıştır.  $OR=9,01$  olduğu durumda Begg testinin performansı özellikle küçük örneklem hacimlerinde düşmüş, Egger testi ise yine her durumda %80 güce ulaşmıştır. Meta analizine alınan çalışma sayısı  $k=20$ 'den  $k=25$ 'e çıktığında bu tez çalışmasında incelenen bütün testlerinin yanlılığı tespit etme gücü bakımından performansları artmıştır.

Simülasyon sonucunda (Tablo 8-12) örneklem hacmi ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=50$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Begg, Egger ve Thompson testleri göstermiştir. Begg, Egger ve Thompson testlerini, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri sadece  $OR=9,01$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Bu tez çalışmasında incelenen hiçbir test  $OR=1,00$  olduğu durumda %80 güce ulaşamamıştır. Begg, Egger ve Thompson testinin yanlılığı tespit etme gücü  $OR=1,00$  değeri hariç her durumda %80'e ulaşmıştır. Meta analizine alınan çalışma sayısı  $k=25$ 'ten  $k=50$ 'ye çıktığında bu tez çalışmasında incelenen bütün testlerinin yanlılığı tespit etme gücü bakımından performansları artmıştır.

Simülasyon sonucunda (Tablo 8-12) örneklem hacmi ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=100$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Egger ve Thompson testleri her koşulda tam olarak %80 güce ulaşarak en iyi performansı göstermiştir. Egger ve Thompson testlerini, yakın bir performansla Begg testi ve daha sonra Harbord ve Schwarzer testleri takip etmektedir. Begg testinin yanlılığı tespit etme gücü sadece  $OR=1,00$  ve örneklem hacmi büyük olduğu durumlarda %80'in altına düşmüştür.

Schwarzer testi sadece  $OR=9,01$  olduğu durum ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Harbord testi ise  $OR \geq 4,02$  olduğu durumlar ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Meta analizine alınan çalışma sayısı  $k=50$ 'den  $k=100$ 'e çıktığında bu tez çalışmasında incelenen bütün testlerinin yanlılığı tespit etme gücü bakımından performansları artmıştır.

Simülasyon sonucunda (Tablo 8-12) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=150$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Begg, Egger ve Thompson testleri her koşulda tam olarak %80 güce ulaşarak en iyi performansı göstermiştir. Begg, Egger ve Thompson testlerini, Harbord ve Schwarzer testleri takip etmektedir. Harbord ve Schwarzer testlerinde OR değeri büyüdükçe ve örneklem hacmi küçüldükçe yanlılığı tespit etme güçleri yükselmekte ve %80 güce ulaşabilmektedir. Meta analizine alınan çalışma sayısı  $k=100$ 'den  $k=150$ 'ye çıktığında bu tez çalışmasında incelenen Begg, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları artmış, Egger ve Thompson, testlerinin ise değişmemiştir.

Simülasyon sonucunda (Tablo 8-12) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve örneklem hacmi  $n=8$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir.

Simülasyon sonucunda (Tablo 8-12) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve örneklem hacmi  $n=10$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Begg ve Egger testleri göstermiştir. Begg ve Egger testlerini, Thompson, Harbord ve Schwarzer testleri takip etmektedir. Genel olarak örnekleme alınan birim sayısının  $n=8$ 'den,  $n=10$ 'a

çıkmasıyla Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselirken, Harbord ve Schwarzer testlerinin düşmüştür.

Simülasyon sonucunda (Tablo 8-12) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve örneklem hacmi  $n=12$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Genel olarak örnekleme alınan birim sayısının  $n=10$ 'dan,  $n=12$ 'ye çıkmasıyla Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Harbord ve Begg testlerinin performansları düşmüş, Schwarzer testinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 8-12) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve örneklem hacmi  $n=15$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Genel olarak örnekleme alınan birim sayısının  $n=12$ 'den,  $n=15$ 'e çıkmasıyla Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Schwarzer testinin performansı düşmüş, Harbord testinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 8-12) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve örneklem hacmi  $n=20$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson ve Harbord testleri takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren test Schwarzer testidir. Schwarzer testi hiçbir durumda %80 güce ulaşamamıştır. Genel olarak örnekleme alınan birim sayısının  $n=15$ 'ten,  $n=20$ 'ye

çıkmasıyla Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Harbord ve Schwarzer testlerinin performansları düşmüş, Begg testinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 8-12) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve örneklem hacmi  $n=25$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson ve Harbord testleri takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren test Schwarzer testidir. Schwarzer testi hiçbir durumda %80 güce ulaşamamıştır. Genel olarak örnekleme alınan birim sayısının  $n=20$ 'den,  $n=25$ 'e çıkmasıyla Begg testinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Harbord testinin performansı düşmüş, Egger ve Thompson testlerinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 8-12) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %10 yanlılık oluşturulduğu ve örneklem hacmi  $n=50$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson testi takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren, hiçbir durumda %80 güce ulaşamayan Harbord ve Schwarzer testleridir. Genel olarak örnekleme alınan birim sayısının  $n=25$ 'ten,  $n=50$ 'ye çıkmasıyla Thompson testinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Harbord testinin performansı düşmüş, Egger ve Begg testlerinin performansı ise değişmemiştir.

Meta analizinde heterojenliğin olmadığı ve meta analizinde %10 yanlılık oluşturulduğu durum dikkate alınarak yapılan simülasyon senaryoları (Tablo 8-12) incelendiğinde örneklem hacmi, meta analizine alınan çalışma sayısı ve OR değeri ayrımı yapılmaksızın genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bütün testlerin her

koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. OR değerinin artmasıyla Egger testinin yanlılığı tespit etme gücü bakımından performansı da yükselmiştir. Begg ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları OR=4,02 değerine kadar yükselmiş, bu OR değerinden itibaren düşmeye başlamıştır. Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları, OR değerinin artmasıyla özellikle OR=4,02 değerinden sonra, örneklem hacminin azalması ve meta analizine alınan çalışma sayılarının artması ile yükselmiştir. Bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları meta analizine alınan çalışma sayıları arttıkça yükselmiştir. Meta analizine alınan çalışma sayısı az iken bu tez çalışmasında kullanılan testlerin hiçbiri yanlılığı tespit etme gücü bakımından iyi performans göstermezken, meta analizine alınan çalışma sayısı yüksek ise Begg, Egger ve Thompson testleri çok iyi performans göstermiştir. Örneklem hacmi arttıkça Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselme eğiliminde iken, Schwarzer ve Harbord testlerinin performansları düşme eğilimindedir.

### **4.3. Meta Analizde %20 Yanlılık Oluşturduğu Duruma Ait Bulgular**

Meta analizine alınan çalışmaların OR değerleri sıralanıp en büyük OR değerlerine sahip olanların %20'si meta analizinden çıkarılarak funnel grafiğinin sağ tarafında asimetri sağlanması ile yanlılık oluşturulduğu durum dikkate alınarak yapılan simülasyon çalışması ile elde edilen veriler analiz edilmiş ve analizlerden elde edilen bulgular tablolar (Tablo 13 – 17) aracılığıyla sunulmuştur. Hasta grubunda pozitif sonuç oranlarına göre ayrı ayrı sunulan tablolarda anakütle OR değerleri,  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  olan örneklem hacimleri ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  olan meta analizine alınan çalışma sayılarına göre funnel grafiği asimetrisinin değerlendirilmesi ile meta analizinde yanlılık olup olmadığına karar vermemizi sağlayan Begg, Egger, Schwarzer, Harbord ve Thompson testlerine ait yanlılığı tespit

etme güçleri verilmiştir. Nominal değer yani Tip-I hata düzeyi  $\alpha=0,10$  olarak alınmıştır.

Meta analizinde %20 yanlılık oluşturulduğu,  $OR=1,00$ ,  $P(S^+ \setminus H^-)=0,50$  ve  $P(S^+ \setminus H^+)=0,50$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 13'te gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayırımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger, Begg ve Thompson testleri göstermiştir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücünün meta analizine alınan çalışma sayısı arttıkça yükseldiği özellikle  $k \geq 50$  olduğu durumlarda %80'nin üzerine çıktığı görülmektedir. Harbord ve Schwarzer testleri her koşulda %80 gücün üzerine çıkamadıklarından yanlılığı tespit etme gücü bakımından performanslarının kötü olduğu söylenebilir. Harbord ve Schwarzer testlerinin gücü meta analizine alınan çalışma sayısının artışıyla az da olsa etkilenecek yükselmiştir. Begg, Egger, Thompson, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü örneklem hacminin değişiminden etkilenmemiştir. Az sayıda çalışma içeren meta analizler için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür.

**Tablo 13.** Meta analizinde %20 yanlışlık oluşturulduğu durumda OR=1,00  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,50$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,416	0,436	0,114	0,080	0,107
	10	0,406	0,444	0,138	0,086	0,123
	12	0,399	0,412	0,129	0,095	0,089
	15	0,403	0,431	0,104	0,096	0,125
	20	0,344	0,390	0,108	0,088	0,108
	25	0,367	0,412	0,100	0,076	0,101
10	50	0,320	0,386	0,100	0,104	0,108
	8	0,412	0,466	0,126	0,083	0,166
	10	0,403	0,458	0,123	0,087	0,177
	12	0,382	0,442	0,155	0,113	0,184
	15	0,400	0,442	0,111	0,100	0,190
	20	0,374	0,428	0,125	0,105	0,176
15	25	0,378	0,421	0,108	0,096	0,167
	50	0,322	0,403	0,104	0,097	0,157
	8	0,487	0,562	0,147	0,077	0,322
	10	0,500	0,568	0,151	0,079	0,338
	12	0,493	0,527	0,148	0,100	0,318
	15	0,507	0,551	0,136	0,100	0,347
20	20	0,482	0,549	0,132	0,102	0,312
	25	0,482	0,538	0,119	0,092	0,309
	50	0,432	0,513	0,101	0,111	0,300
	8	0,572	0,645	0,144	0,072	0,451
	10	0,547	0,645	0,125	0,074	0,464
	12	0,576	0,636	0,143	0,093	0,455
25	15	0,607	0,666	0,130	0,095	0,466
	20	0,600	0,660	0,150	0,107	0,458
	25	0,578	0,625	0,125	0,109	0,419
	50	0,509	0,596	0,113	0,110	0,413
	8	0,649	0,712	0,157	0,067	0,533
	10	0,661	0,742	0,157	0,081	0,573
50	12	0,620	0,715	0,160	0,091	0,536
	15	0,688	0,741	0,147	0,096	0,590
	20	0,646	0,726	0,139	0,107	0,547
	25	0,630	0,706	0,122	0,108	0,518
	50	0,591	0,670	0,120	0,119	0,488
	8	0,866	0,892	0,183	0,065	0,807
100	10	0,883	0,936	0,202	0,100	0,868
	12	0,863	0,924	0,189	0,085	0,843
	15	0,893	0,922	0,172	0,103	0,854
	20	0,863	0,905	0,170	0,101	0,808
	25	0,856	0,903	0,147	0,101	0,809
	50	0,840	0,902	0,149	0,103	0,800
150	8	0,990	0,996	0,284	0,070	0,979
	10	0,988	0,997	0,300	0,100	0,991
	12	0,983	0,994	0,234	0,102	0,989
	15	0,990	0,994	0,250	0,127	0,982
	20	0,985	0,989	0,209	0,104	0,977
	25	0,982	0,986	0,225	0,111	0,974
150	50	0,976	0,988	0,164	0,095	0,976
	8	0,998	1,000	0,382	0,064	1,000
	10	0,999	1,000	0,379	0,117	1,000
	12	0,997	0,999	0,254	0,123	0,994
	15	1,000	1,000	0,266	0,112	0,997
	20	0,996	0,997	0,270	0,115	0,995
150	25	0,997	1,000	0,278	0,126	0,998
	50	0,996	1,000	0,192	0,110	1,000

Meta analizinde %20 yanlılık oluşturulduğu,  $OR=1,50$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,60$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 14’de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Begg testi göstermiştir. Begg testini, bu teste çok yakın bir performansla Egger testi ve daha sonra Thompson testi takip etmektedir. Harbord testi %50 gücün üzerine çıkamadığından en az güce sahip olduğu söylenebilir. Schwarzer testinin yanlılık belirleme gücü meta analizine alınan çalışma sayısı fazla ise (50, 100, 150 gibi) az da olsa yükselmeye başlamış, meta analizine alınan çalışma sayısı  $k=150$  olduğu durumda ancak %50’nin üzerine çıkabilmiştir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücünün meta analizine alınan çalışma sayısı arttıkça yükseldiği, Begg ve Egger testlerinin özellikle  $k \geq 15$ , Thompson testinin ise  $k \geq 20$  olduğu durumlarda güçlerinin %80’nin üzerine çıktığı görülmektedir. Harbord testinin gücünde ise meta analizine alınan çalışma sayısına göre bir değişim gerçekleşmemiştir. Az sayıda çalışma içeren meta analizler için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı örnekleme alınan birim sayısı arttıkça yükselme eğiliminde iken, Harbord ve Schwarzer testlerinin gücü örneklem hacminin değişiminden etkilenmemiştir. OR değerinin artışıyla birlikte genel olarak Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Harbord ve Schwarzer testlerinin gücünde bir değişim meydana gelmemiştir.



**Tablo 14.** Meta analizinde %20 yanlışlık oluşturulduğu durumda  $OR=1,50$   $P(S^+|H^+)=0,50$  ve  $P(S^+|H^+)=0,60$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,658	0,630	0,097	0,062	0,159
	10	0,670	0,658	0,120	0,090	0,196
	12	0,667	0,669	0,112	0,093	0,174
	15	0,696	0,656	0,109	0,092	0,165
	20	0,718	0,671	0,131	0,106	0,196
	25	0,688	0,666	0,099	0,111	0,178
10	50	0,665	0,667	0,100	0,101	0,205
	8	0,695	0,666	0,113	0,080	0,259
	10	0,774	0,712	0,120	0,082	0,332
	12	0,776	0,756	0,113	0,076	0,340
	15	0,782	0,718	0,098	0,102	0,334
	20	0,792	0,730	0,123	0,100	0,349
15	25	0,816	0,753	0,096	0,091	0,330
	50	0,808	0,762	0,091	0,089	0,381
	8	0,870	0,826	0,109	0,063	0,549
	10	0,913	0,862	0,140	0,072	0,607
	12	0,916	0,872	0,125	0,074	0,637
	15	0,932	0,876	0,126	0,088	0,617
20	20	0,943	0,872	0,140	0,099	0,617
	25	0,950	0,884	0,133	0,087	0,636
	50	0,841	0,885	0,128	0,101	0,627
	8	0,936	0,899	0,116	0,059	0,697
	10	0,965	0,935	0,133	0,057	0,772
	12	0,965	0,926	0,158	0,085	0,793
25	15	0,987	0,954	0,155	0,097	0,797
	20	0,985	0,936	0,157	0,100	0,775
	25	0,991	0,949	0,136	0,097	0,804
	50	0,984	0,945	0,125	0,107	0,793
	8	0,983	0,948	0,116	0,062	0,816
	10	0,985	0,971	0,140	0,057	0,865
50	12	0,994	0,973	0,156	0,077	0,889
	15	0,996	0,979	0,157	0,097	0,890
	20	0,996	0,977	0,151	0,099	0,883
	25	0,996	0,978	0,128	0,102	0,885
	50	0,992	0,982	0,111	0,097	0,888
	8	1,000	0,999	0,164	0,048	0,992
100	10	1,000	0,999	0,234	0,067	0,992
	12	1,000	1,000	0,234	0,093	0,994
	15	1,000	1,000	0,257	0,118	0,998
	20	1,000	1,000	0,230	0,091	0,990
	25	1,000	0,998	0,195	0,105	0,992
	50	1,000	0,998	0,119	0,116	0,993
150	8	1,000	1,000	0,239	0,041	1,000
	10	1,000	1,000	0,363	0,078	1,000
	12	1,000	1,000	0,436	0,093	1,000
	15	1,000	1,000	0,438	0,132	1,000
	20	1,000	1,000	0,371	0,105	1,000
	25	1,000	1,000	0,329	0,118	1,000
150	50	1,000	1,000	0,153	0,092	0,999
	8	1,000	1,000	0,342	0,053	1,000
	10	1,000	1,000	0,505	0,092	1,000
	12	1,000	1,000	0,568	0,145	1,000
	15	1,000	1,000	0,565	0,162	1,000
	20	1,000	1,000	0,493	0,154	1,000
150	25	1,000	1,000	0,459	0,134	1,000
	50	1,000	1,000	0,197	0,098	1,000

Meta analizinde %20 yanlılık oluşturulduğu,  $OR=2,34$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,70$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 15'te gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayırımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, bu teste çok yakın bir performansla Begg testi ve daha sonra Thompson testi takip etmektedir. Az sayıda çalışma içeren meta analizler için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür. Begg ve Egger testleri  $k \geq 10$ , Thompson testi ise  $k \geq 20$  olduğu durumlarda %80'nin üzerinde güç değerleri elde etmiştir. Schwarzer ve Harbord testleri en az güce sahiptir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısı arttıkça yükselmiştir. Schwarzer testinin gücü meta analizine alınan çalışma sayısı  $k \geq 50$  durumlarında az da olsa bir artış göstermiştir. Meta analizine alınan çalışma sayısı arttıkça Harbord testinin gücü yalnızca örnekleme alınan birim sayısı  $n=8$  olduğu durumda yükselmiştir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı örnekleme alınan birim sayısı arttıkça yükselme eğiliminde iken, Harbord ve Schwarzer testlerinin gücü örneklem hacminin değişiminden etkilenmemiştir. OR değerinin artışıyla birlikte Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiştir. OR değerinin artışıyla birlikte Begg, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı değişmemiştir.

**Tablo 15.** Meta analizinde %20 yanlışlık oluşturulduğu durumda  $OR=2,34$   $P(S^+|H^+)=0,50$  ve  $P(S^+|H^+)=0,70$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,674	0,695	0,103	0,056	0,147
	10	0,711	0,729	0,084	0,062	0,178
	12	0,700	0,720	0,105	0,085	0,166
	15	0,679	0,723	0,103	0,089	0,190
	20	0,682	0,720	0,103	0,091	0,215
	25	0,653	0,707	0,128	0,113	0,196
10	50	0,589	0,686	0,105	0,105	0,202
	8	0,783	0,774	0,093	0,061	0,292
	10	0,835	0,796	0,084	0,056	0,374
	12	0,801	0,809	0,089	0,068	0,348
	15	0,787	0,814	0,110	0,097	0,379
	20	0,779	0,802	0,108	0,103	0,366
15	25	0,780	0,835	0,103	0,102	0,378
	50	0,730	0,803	0,101	0,105	0,409
	8	0,937	0,900	0,090	0,056	0,580
	10	0,966	0,939	0,072	0,050	0,656
	12	0,936	0,930	0,100	0,057	0,663
	15	0,936	0,936	0,107	0,091	0,654
20	20	0,942	0,926	0,117	0,099	0,688
	25	0,937	0,939	0,100	0,096	0,683
	50	0,905	0,929	0,118	0,113	0,689
	8	0,972	0,957	0,088	0,065	0,741
	10	0,994	0,982	0,083	0,044	0,831
	12	0,985	0,978	0,092	0,075	0,838
25	15	0,982	0,977	0,101	0,080	0,835
	20	0,983	0,962	0,116	0,100	0,831
	25	0,981	0,971	0,100	0,098	0,851
	50	0,973	0,972	0,123	0,123	0,854
	8	0,994	0,988	0,093	0,075	0,850
	10	0,998	0,991	0,078	0,058	0,909
50	12	0,997	0,993	0,094	0,052	0,929
	15	0,997	0,994	0,099	0,081	0,931
	20	0,994	0,989	0,106	0,087	0,916
	25	0,994	0,992	0,100	0,099	0,931
	50	0,988	0,988	0,098	0,094	0,910
	8	1,000	1,000	0,094	0,148	0,996
100	10	1,000	1,000	0,107	0,074	0,999
	12	1,000	1,000	0,121	0,066	0,999
	15	1,000	1,000	0,136	0,077	1,000
	20	1,000	1,000	0,121	0,089	0,999
	25	1,000	1,000	0,118	0,104	1,000
	50	1,000	1,000	0,107	0,101	0,998
150	8	1,000	1,000	0,112	0,342	1,000
	10	1,000	1,000	0,099	0,102	1,000
	12	1,000	1,000	0,125	0,059	1,000
	15	1,000	1,000	0,161	0,083	1,000
	20	1,000	1,000	0,126	0,098	1,000
	25	1,000	1,000	0,118	0,087	1,000
150	50	1,000	1,000	0,105	0,092	1,000
	8	1,000	1,000	0,126	0,535	1,000
	10	1,000	1,000	0,103	0,150	1,000
	12	1,000	1,000	0,135	0,062	1,000
	15	1,000	1,000	0,199	0,060	1,000
	20	1,000	1,000	0,135	0,095	1,000
150	25	1,000	1,000	0,131	0,097	1,000
	50	1,000	1,000	0,117	0,098	1,000

Meta analizinde %20 yanlışlık oluşturulduğu,  $OR=4,02$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,80$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 16'da gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, bu teste çok yakın bir performansla Begg testi, Thompson testi ve daha sonra sırası ile Harbord ve Schwarzer testleri takip etmektedir. Az sayıda çalışma içeren meta analizler için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür. Begg, Egger, Thompson, Schwarzer ve Harbord testlerinin yanlışlığı tespit etme gücü meta analizine alınan çalışma sayısı arttıkça yükselmiştir. Harbord ve Schwarzer testlerinde örneklemdeki birim sayısının azalması ile güç yükselmiştir. Meta analizine alınan çalışma sayısı  $k \geq 100$  gibi büyük ve örnekleme alınan birim sayısı  $n \leq 10$  gibi küçük değerler aldıkça Harbord ve Schwarzer testlerinin güç bakımından iyi performans gösterdiği söylenebilir. Begg, Egger ve Thompson testlerinin yanlışlığı tespit etme gücü bakımından performansı örnekleme alınan birim sayısı arttıkça yükselme eğilimindedir. OR değerinin artışıyla birlikte Schwarzer ve Harbord testlerinin yanlışlığı tespit etme gücü bakımından performansı yükselmiş, Begg ve Thompson testlerinin yanlışlığı tespit etme gücü bakımından performansı düşmüş, Egger testinin performansı ise değişmemiştir.

**Tablo 16.** Meta analizinde %20 yanlışlık oluşturulduğu durumda  $OR=4,02$   $P(S^+|H^+)=0,50$  ve  $P(S^+|H^+)=0,80$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,530	0,628	0,185	0,133	0,070
	10	0,621	0,689	0,091	0,073	0,097
	12	0,612	0,699	0,102	0,082	0,110
	15	0,594	0,728	0,102	0,079	0,143
	20	0,661	0,780	0,105	0,090	0,183
	25	0,642	0,778	0,097	0,079	0,198
10	8	0,627	0,704	0,200	0,161	0,187
	10	0,726	0,785	0,110	0,098	0,238
	12	0,727	0,809	0,093	0,076	0,234
	15	0,720	0,844	0,110	0,067	0,325
	20	0,793	0,874	0,108	0,075	0,392
	25	0,800	0,884	0,096	0,087	0,437
15	8	0,765	0,850	0,272	0,251	0,403
	10	0,875	0,927	0,127	0,130	0,500
	12	0,885	0,938	0,097	0,094	0,527
	15	0,919	0,959	0,093	0,073	0,632
	20	0,942	0,979	0,109	0,073	0,726
	25	0,938	0,974	0,101	0,093	0,740
20	8	0,917	0,957	0,102	0,090	0,733
	10	0,876	0,945	0,329	0,327	0,581
	12	0,944	0,976	0,161	0,201	0,689
	15	0,958	0,988	0,107	0,113	0,753
	20	0,970	0,988	0,089	0,068	0,844
	25	0,984	0,995	0,111	0,078	0,879
25	8	0,986	0,992	0,099	0,067	0,894
	10	0,973	0,988	0,089	0,089	0,895
	12	0,930	0,981	0,370	0,416	0,735
	15	0,983	0,992	0,169	0,238	0,848
	20	0,986	0,994	0,112	0,121	0,877
	25	0,994	0,997	0,082	0,070	0,941
50	8	0,999	1,000	0,109	0,088	0,950
	10	0,996	0,999	0,086	0,058	0,968
	12	0,990	0,997	0,100	0,098	0,939
	15	0,997	1,000	0,583	0,725	0,988
	20	1,000	1,000	0,312	0,511	0,993
	25	1,000	1,000	0,173	0,286	0,998
100	8	1,000	1,000	0,097	0,108	1,000
	10	1,000	1,000	0,106	0,072	1,000
	12	1,000	1,000	0,114	0,071	1,000
	15	1,000	1,000	0,102	0,094	1,000
	20	1,000	1,000	0,831	0,954	1,000
	25	1,000	1,000	0,554	0,869	1,000
150	8	1,000	1,000	0,259	0,528	1,000
	10	1,000	1,000	0,129	0,223	1,000
	12	1,000	1,000	0,108	0,089	1,000
	15	1,000	1,000	0,124	0,070	1,000
	20	1,000	1,000	0,102	0,081	1,000
	25	1,000	1,000	0,922	0,991	1,000
150	8	1,000	1,000	0,746	0,960	1,000
	10	1,000	1,000	0,321	0,685	1,000
	12	1,000	1,000	0,146	0,317	1,000
	15	1,000	1,000	0,133	0,117	1,000
	20	1,000	1,000	0,160	0,087	1,000
	25	1,000	1,000	0,140	0,090	1,000

Meta analizinde %20 yanlılık oluşturulduğu,  $OR=9,01$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,90$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 17’de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, Begg, Thompson, Harbord ve Schwarzer testleri takip etmektedir. Bu tez çalışmasında incelenen bütün testlerin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısının artmasıyla yükselmiştir. Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları meta analizine alınan çalışma sayısı arttıkça ve örnekleme alınan birim sayısı azaldıkça yükselmiştir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları örnekleme alınan birim sayısı arttıkça yükselmiştir. OR değerinin 4,02’den 9,01’ye artmasıyla birlikte genel olarak bakıldığında Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları düşmüş, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

**Tablo 17.** Meta analizinde %20 yanlışlık oluşturulduğu durumda OR=9,01  $P(S^+|H^+)=0,50$  ve  $P(S^+|H^+)=0,90$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,428	0,523	0,547	0,501	0,018
	10	0,492	0,538	0,378	0,413	0,033
	12	0,464	0,507	0,296	0,356	0,017
	15	0,511	0,641	0,151	0,210	0,038
	20	0,617	0,749	0,104	0,162	0,076
	50	0,777	0,889	0,079	0,094	0,261
10	8	0,462	0,596	0,611	0,579	0,055
	10	0,579	0,640	0,431	0,461	0,083
	12	0,545	0,649	0,330	0,379	0,074
	15	0,591	0,777	0,200	0,255	0,119
	20	0,753	0,885	0,115	0,168	0,199
	50	0,819	0,934	0,083	0,108	0,284
15	8	0,549	0,721	0,783	0,783	0,157
	10	0,676	0,804	0,579	0,639	0,197
	12	0,652	0,794	0,442	0,569	0,221
	15	0,754	0,919	0,236	0,350	0,322
	20	0,906	0,972	0,134	0,235	0,498
	50	0,954	0,986	0,096	0,153	0,630
20	8	0,637	0,843	0,863	0,880	0,273
	10	0,765	0,898	0,673	0,788	0,343
	12	0,755	0,910	0,548	0,696	0,338
	15	0,852	0,972	0,320	0,480	0,515
	20	0,971	0,992	0,150	0,324	0,706
	50	0,985	0,999	0,088	0,212	0,856
25	8	0,680	0,907	0,928	0,945	0,372
	10	0,829	0,945	0,736	0,864	0,472
	12	0,830	0,944	0,646	0,807	0,487
	15	0,913	0,990	0,368	0,587	0,698
	20	0,990	1,000	0,172	0,412	0,872
	50	0,998	0,999	0,098	0,252	0,952
50	8	0,998	1,000	0,114	0,127	0,986
	8	0,866	0,992	0,994	0,998	0,797
	10	0,968	0,999	0,954	0,988	0,904
	12	0,968	1,000	0,886	0,976	0,904
	15	0,997	1,000	0,601	0,891	0,984
	20	1,000	1,000	0,265	0,681	1,000
100	25	1,000	1,000	0,115	0,489	1,000
	50	1,000	1,000	0,110	0,196	1,000
	8	0,977	1,000	1,000	1,000	0,991
	10	0,999	1,000	1,000	1,000	1,000
	12	0,998	1,000	0,992	0,999	0,998
	15	1,000	1,000	0,889	0,994	1,000
150	20	1,000	1,000	0,428	0,928	1,000
	25	1,000	1,000	0,173	0,772	1,000
	50	1,000	1,000	0,117	0,288	1,000
	8	0,994	1,000	1,000	1,000	1,000
	10	1,000	1,000	1,000	1,000	0,997
	12	0,999	1,000	1,000	1,000	1,000
150	15	1,000	1,000	0,997	1,000	1,000
	20	1,000	1,000	0,591	0,979	1,000
	25	1,000	1,000	0,192	0,916	1,000
	50	1,000	1,000	0,148	0,405	1,000

Simülasyon sonucunda (Tablo 13–17) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=8$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Begg, Schwarzer, Harbord ve Thompson testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger testi ise sadece  $OR= 9,01$  ve  $n \geq 25$  olduğu durumlarda %80 güce ulaşabilmiştir.

Simülasyon sonucunda (Tablo 13–17) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=10$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi takip etmektedir. Schwarzer, Harbord ve Thompson testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger testi  $OR \geq 2,34$  ve örneklem hacmi büyük olduğu durumlarda %80 güce ulaşabilmiştir. Meta analizine alınan çalışma sayısı  $k=8$ 'den  $k=10$ 'a çıktığında Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 13–17) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=15$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra çok da iyi olmayan bir performansla Thompson testi takip etmektedir. Schwarzer ve Harbord testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger ve Begg testi  $OR= 1,00$  olduğu durum hariç %80 güce ulaşmış, Thompson testi ise sadece  $OR=9,01$  ve  $n=50$  olduğu durumda %80 güce ulaşabilmiştir. Begg testinin yanlılığı tespit etme gücü performansı  $OR=2,34$  değerine kadar yükselmiş, bu OR değerinden sonra özellikle küçük örneklem hacimlerinde düşmeye başlamıştır.  $OR=1,50$  ve  $2,34$  iken Begg testi, Egger testi ile aynı



performansı göstermiş,  $OR \geq 4,02$  olduğu durumlarda Egger testinden daha kötü performans göstermiştir. Egger testinin yanlılığı tespit etme gücü performansı  $OR=9,01$  değerinde özellikle küçük örneklem hacimlerinde düşmeye başlamıştır. Meta analizine alınan çalışma sayısı  $k=10$ 'dan  $k=15$ 'e çıktığında Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 13–17) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=20$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Schwarzer ve Harbord testleri takip etmektedir. Schwarzer ve Harbord testleri sadece  $OR=9,01$  ve  $n=8$  olduğu durumda %80 güce ulaşabilmiştir. Bu tez çalışmasında incelenen hiçbir test  $OR=1,00$  olduğu durumda %80 güce ulaşamamıştır. Thompson testinin yanlılığı tespit etme gücü bakımından performansı  $OR=2,34$  değerine kadar yükselmiş, daha sonra OR değerinin artışı ile birlikte küçük örneklem hacimlerinde düşmeye başlamıştır. Begg testinin yanlılığı tespit etme gücü  $OR=1,50$ ,  $OR=2,34$  ve  $OR=4,02$  olduğu her durumda %80'e ulaşmış,  $OR=9,01$  olduğu durumda performansı özellikle küçük örneklem hacimlerinde düşmüştür. Egger testinin yanlılığı tespit etme gücü bakımından performansı  $OR=1,00$  hariç her durumda %80'e ulaşmıştır. Meta analizine alınan çalışma sayısı  $k=15$ 'ten  $k=20$ 'ye çıktığında bu tez çalışmasında incelenen bütün testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 13–17) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=25$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri sadece  $OR=9,01$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Bu tez çalışmasında

incelenen hiçbir test  $OR=1,00$  olduğu durumda %80 güce ulaşamamıştır. Thompson testinin yanlılığı tespit etme gücü bakımından performansı  $OR=2,34$  değerine kadar yükselmiş, daha sonra  $OR$  değerinin artışı ile birlikte performansı küçük örneklem hacimlerinde düşmeye başlamıştır. Begg ve Egger testlerinin yanlılığı tespit etme gücü  $OR=1,50$ ,  $OR=2,34$  ve  $OR=4,02$  olduğu her durumda %80'e ulaşmıştır.  $OR=9,01$  olduğu durumda Begg testinin performansı özellikle küçük örneklem hacimlerinde düşmüş, Egger testi ise yine her durumda %80 güce ulaşmıştır. Meta analizine alınan çalışma sayısı  $k=20$ 'den  $k=25$ 'e çıktığında bu Begg, Harbord ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Egger ve Schwarzer testlerinin performansları değişmemiştir.

Simülasyon sonucunda (Tablo 13–17) örneklem hacmi ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=50$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Begg ve Egger testleri her koşulda tam olarak %80 güce ulaşarak en iyi performansı göstermiştir. Begg ve Egger testlerini, çok yakın bir performansla Thompson testi ve daha sonra Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri sadece  $OR=9,01$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Thompson testinin yanlılığı tespit etme gücü  $OR=9,01$  ve  $n=8$  olduğu durum hariç her durumda %80'e ulaşmıştır. Meta analizine alınan çalışma sayısı  $k=25$ 'ten  $k=50$ 'ye çıktığında bu tez çalışmasında incelenen bütün testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 13–17) örneklem hacmi ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=100$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Begg, Egger ve Thompson testleri her koşulda tam olarak %80 güce ulaşarak en iyi performansı göstermiştir. Begg, Egger ve Thompson testlerini, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri  $OR \geq 4,02$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Meta analizine alınan çalışma sayısı  $k=50$ 'den  $k=100$ 'e çıktığında bu tez çalışmasında incelenen Thompson, Harbord ve Schwarzer testlerinin yanlılığı

tespit etme gücü bakımından performansları yükselmiş, Begg ve Egger testlerinin ise değişmemiştir.

Simülasyon sonucunda (Tablo 13–17) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=150$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Begg, Egger ve Thompson testleri her koşulda tam olarak %80 güce ulaşarak en iyi göstermiştir. Begg, Egger ve Thompson testlerini, Harbord ve Schwarzer testleri takip etmektedir. Harbord ve Schwarzer testlerinde OR değeri arttıkça ve örneklem hacmi azaldıkça yanlılığı tespit etme güçleri yükselmekte ve %80 güce ulaşabilmektedir. Meta analizine alınan çalışma sayısı  $k=100$ 'den  $k=150$ 'ye çıktığında bu tez çalışmasında incelenen Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Thompson, Begg ve Egger testlerinin ise değişmemiştir.

Simülasyon sonucunda (Tablo 13–17) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve örneklem hacmi  $n=8$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=4,02$  değerine kadar yükseliş eğiliminde iken bu değerden sonra düşmeye başlamıştır. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=4,02$  değerinden sonra yükselmiştir.

Simülasyon sonucunda (Tablo 13–17) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve örneklem hacmi  $n=10$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg ve Thompson testlerinin yanlılığı tespit

etme gücü bakımından performansları  $OR=4,02$  değerine kadar yükseliş eğiliminde iken bu değerden sonra düşmeye başlamıştır. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=4,02$  değerinden sonra yükselmiştir. Egger testinin yanlılığı tespit etme gücü bakımından performansı ise  $OR$  değeri 1,00 değerinden 1,50 değerine artarken yükselmiş, daha sonra  $OR$  değerinin artışı ile değişmemiştir. Genel olarak örnekleme alınan birim sayısının  $n=8$ 'den,  $n=10$ 'a çıkmasıyla Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselirken, Harbord ve Schwarzer testlerinin düşmüştür.

Simülasyon sonucunda (Tablo 13–17) meta analizine alınan çalışma sayısı ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve örneklem hacmi  $n=12$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=4,02$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren düşmeye başlamıştır. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=4,02$  değerinden itibaren yükselmiştir. Egger testinin yanlılığı tespit etme gücü bakımından performansı ise  $OR=4,02$  değerine kadar yükseliş eğiliminde iken,  $OR=9,01$  değerinden itibaren düşmeye başlamıştır. Genel olarak örnekleme alınan birim sayısının  $n=10$ 'dan,  $n=12$ 'ye çıkmasıyla Egger testinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Harbord testinin performansı düşmüş, Begg, Schwarzer ve Thompson testlerinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 13–17) meta analizine alınan çalışma sayısı ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve örneklem hacmi  $n=15$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=9,01$  değerine kadar yükseliş

eğiliminde iken bu değerden itibaren düşmeye başlamıştır. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=9,01$  değerinden itibaren yükselmiştir. Genel olarak örnekleme alınan birim sayısının  $n=12$ 'den,  $n=15$ 'e çıkmasıyla Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Schwarzer ve Harbord testlerinin performansları düşmüş, Begg testinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 13–17) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve örneklem hacmi  $n=20$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson ve Harbord testleri takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren test Schwarzer testidir. Schwarzer testi hiçbir durumda %80 güce ulaşamamıştır. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=9,01$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren Thompson testinin performansı düşmeye başlamıştır. Harbord testinin yanlılığı tespit etme gücü bakımından performansı ise  $OR=9,01$  değerinden itibaren yükselmiştir. Genel olarak örnekleme alınan birim sayısının  $n=15$ 'ten,  $n=20$ 'ye çıkmasıyla Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Harbord ve Schwarzer testlerinin performansları düşmüştür.

Simülasyon sonucunda (Tablo 13–17) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve örneklem hacmi  $n=25$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson ve Harbord testleri takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren test Schwarzer testidir. Schwarzer testi hiçbir durumda %80 güce ulaşamamıştır. OR değeri arttıkça Begg, Egger, Thompson ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğilimindedir. Egger

testi  $OR=9,01$  olduğu her durumda %80 güce ulaşmıştır. Genel olarak örnekleme alınan birim sayısının  $n=20$ 'den,  $n=25$ 'e çıkmasıyla Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Harbord testinin performansı düşmüş, Schwarzer testinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 13–17) meta analizine alınan çalışma sayısı ve OR değerleri ayırımı yapılmaksızın meta analizinde %20 yanlılık oluşturulduğu ve örneklem hacmi  $n=50$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson testi takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren, hiçbir durumda %80 güce ulaşamayan testler Harbord ve Schwarzer testleridir. OR değeri arttıkça Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğilimindedir. Egger testi  $OR=9,01$  olduğu her durumda %80 güce ulaşmıştır. Genel olarak örnekleme alınan birim sayısının  $n=25$ 'ten,  $n=50$ 'ye çıkmasıyla Harbord ve Begg testlerinin yanlılığı tespit etme gücü bakımından performansı düşmüş, Thompson, Egger ve Schwarzer testlerinin performansı ise değişmemiştir.

Meta analizinde %20 yanlılık oluşturulduğu durum dikkate alınarak yapılan simülasyon senaryoları (Tablo 13–17) incelendiğinde örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayırımı yapılmaksızın genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. OR değerinin artmasıyla Egger testinin yanlılığı tespit etme gücü bakımından performansı  $OR=9,01$  değerine kadar yükselmiş, bu değerden itibaren düşmüştür. Begg ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=4,02$  değerine kadar yükselmiş, bu OR değerinden itibaren düşmeye başlamıştır. Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları, OR değerinin artmasıyla özellikle  $OR=4,02$  değerinden sonra, örneklem hacminin azalması ve meta analizine alınan çalışma

sayılarının artması ile yükselmiştir. Bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları meta analizine alınan çalışma sayıları arttıkça yükselmiştir. Meta analizine alınan çalışma sayısı az iken bu tez çalışmasında kullanılan testlerin hiçbiri yanlılığı tespit etme gücü bakımından iyi performans göstermezken, meta analizine alınan çalışma sayısı fazla ise Begg, Egger ve Thompson testleri çok iyi performans göstermiştir. Örneklem hacmi arttıkça Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselme eğiliminde iken, Schwarzer ve Harbord testlerinin performansları düşme eğilimindedir.

#### **4.4. Meta Analizde %40 Yanlılık Oluşturduğu Duruma Ait Bulgular**

Meta analizine alınan çalışmaların OR değerleri sıralanıp en büyük OR değerlerine sahip olanların %40'ı meta analizinden çıkarılarak funnel grafiğinin sağ tarafında asimetri sağlanması ile yanlılık oluşturulduğu durum dikkate alınarak yapılan simülasyon çalışması ile elde edilen veriler analiz edilmiş ve analizlerden elde edilen bulgular tablolar (Tablo 18–22) aracılığıyla sunulmuştur. Hasta grubunda pozitif sonuç oranlarına göre ayrı ayrı sunulan tablolarda anakütle OR değerleri,  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  olan örneklem hacimleri ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  olan meta analizine alınan çalışma sayılarına göre funnel grafiği asimetrisinin değerlendirilmesi ile meta analizinde yanlılık olup olmadığına karar vermemizi sağlayan Begg, Egger, Schwarzer, Harbord ve Thompson testlerine ait yanlılığı tespit etme güçleri verilmiştir. Nominal değer yani Tip-I hata düzeyi  $\alpha=0,10$  olarak alınmıştır.

Meta analizinde %40 yanlılık oluşturulduğu,  $OR=1,00$ ,  $P(S^+ \setminus H^-)=0,50$  ve  $P(S^+ \setminus H^+)=0,50$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 18'de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Begg testi göstermiştir. Begg testini, bu teste çok yakın

bir performans ile Egger testi, daha sonra Thompson testi takip etmektedir. Harbord ve Schwarzer testleri her koşulda %80 gücün üzerine çıkamadıklarından yanlılığı tespit etme gücü bakımından performanslarının kötü olduğu söylenebilir. Az sayıda çalışma içeren meta analizler için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür, %80'e ulaşamamıştır. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı meta analizine alınan çalışma sayısı arttıkça yükselmiştir. Begg testi  $k \geq 20$ , Egger testi  $k \geq 25$ , Thompson testi ise  $k \geq 50$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermişlerdir. Harbord ve Schwarzer testlerinin gücü meta analizine alınan çalışma sayısının artışından etkilenerek yükselse de %80 güce ulaşamamışlardır. Begg, Egger, Thompson, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü örneklem hacminin değişiminden etkilenmemiştir.



**Tablo 18.** Meta analizinde %40 yanlışlık oluşturulduğu durumda OR=1,00  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,50$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,487	0,508	0,126	0,090	0,091
	10	0,463	0,520	0,133	0,095	0,109
	12	0,474	0,487	0,135	0,103	0,088
	15	0,474	0,494	0,130	0,108	0,094
	20	0,445	0,465	0,111	0,088	0,101
	25	0,454	0,472	0,115	0,086	0,097
10	50	0,386	0,458	0,119	0,105	0,088
	8	0,619	0,607	0,139	0,075	0,188
	10	0,605	0,622	0,131	0,103	0,195
	12	0,563	0,589	0,143	0,093	0,182
	15	0,634	0,603	0,146	0,105	0,216
	20	0,597	0,572	0,138	0,097	0,209
15	25	0,590	0,587	0,126	0,103	0,184
	50	0,510	0,532	0,130	0,102	0,172
	8	0,765	0,749	0,162	0,090	0,402
	10	0,755	0,756	0,163	0,087	0,407
	12	0,762	0,743	0,157	0,110	0,401
	15	0,777	0,727	0,165	0,107	0,426
20	20	0,765	0,728	0,146	0,099	0,410
	25	0,781	0,734	0,151	0,098	0,368
	50	0,686	0,668	0,127	0,114	0,358
	8	0,852	0,832	0,171	0,079	0,572
	10	0,857	0,861	0,196	0,103	0,592
	12	0,843	0,850	0,194	0,105	0,596
25	15	0,883	0,847	0,172	0,107	0,576
	20	0,868	0,823	0,162	0,093	0,555
	25	0,873	0,820	0,155	0,109	0,521
	50	0,814	0,786	0,135	0,114	0,516
	8	0,901	0,895	0,180	0,088	0,696
	10	0,921	0,919	0,224	0,114	0,719
50	12	0,913	0,915	0,205	0,094	0,714
	15	0,944	0,912	0,183	0,094	0,714
	20	0,919	0,876	0,187	0,110	0,648
	25	0,925	0,879	0,180	0,113	0,655
	50	0,888	0,857	0,164	0,115	0,618
	8	0,993	0,991	0,268	0,092	0,955
100	10	0,995	0,991	0,330	0,127	0,967
	12	0,993	0,988	0,305	0,123	0,950
	15	0,996	0,993	0,311	0,125	0,953
	20	0,994	0,988	0,309	0,116	0,932
	25	0,995	0,992	0,301	0,123	0,930
	50	0,988	0,976	0,218	0,119	0,910
150	8	1,000	1,000	0,490	0,123	1,000
	10	1,000	1,000	0,473	0,158	1,000
	12	1,000	1,000	0,474	0,168	0,998
	15	1,000	1,000	0,501	0,156	0,998
	20	1,000	1,000	0,515	0,143	0,999
	25	1,000	1,000	0,491	0,140	0,997
150	50	1,000	0,999	0,334	0,106	0,996
	8	1,000	1,000	0,623	0,176	1,000
	10	1,000	1,000	0,580	0,241	1,000
	12	1,000	1,000	0,610	0,199	1,000
	15	1,000	1,000	0,636	0,174	1,000
	20	1,000	1,000	0,646	0,171	1,000
150	25	1,000	1,000	0,647	0,156	1,000
	50	1,000	1,000	0,446	0,123	1,000

Meta analizinde %40 yanlılık oluşturulduğu,  $OR=1,50$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,60$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 19’da gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Begg testi göstermiştir. Begg testini, bu teste çok yakın bir performansla Egger testi ve daha sonra Thompson testi takip etmektedir. Harbord ve Schwarzer testleri her koşulda %80 gücün üzerine çıkamadıklarından yanlılığı tespit etme gücü bakımından performanslarının kötü olduğu söylenebilir. Az sayıda çalışma içeren meta analizler ( $k=8$ ) için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür, %80’e ulaşamamıştır. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı meta analizine alınan çalışma sayısı arttıkça yükselmiştir. Begg testi  $k \geq 10$ , Egger testi  $k \geq 15$ , Thompson testi ise  $k \geq 25$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermişlerdir. Schwarzer testinin gücü meta analizine alınan çalışma sayısının artışıyla etkilenerek yükselse de %80 güce ulaşamamıştır. Harbord testinin gücünde ise meta analizine alınan çalışma sayısına göre bir değişim gerçekleşmemiştir. Bu tez çalışmasında ele alınan bütün testlerin gücünün örneklem hacmine göre değişmediği söylenebilir. OR değerinin artışıyla birlikte genel olarak Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Schwarzer ve Harbord testlerinin performanslarında bir değişim meydana gelmemiştir.

**Tablo 19.** Meta analizinde %40 yanlışlık oluşturulduğu durumda OR=1,50  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,60$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,694	0,666	0,106	0,072	0,131
	10	0,731	0,703	0,116	0,090	0,165
	12	0,702	0,704	0,103	0,070	0,151
	15	0,697	0,679	0,119	0,103	0,145
	20	0,708	0,670	0,114	0,102	0,159
	25	0,695	0,675	0,116	0,092	0,150
10	50	0,676	0,675	0,102	0,096	0,187
	8	0,807	0,775	0,106	0,069	0,230
	10	0,862	0,814	0,128	0,087	0,302
	12	0,834	0,792	0,104	0,088	0,282
	15	0,828	0,778	0,104	0,102	0,271
	20	0,832	0,775	0,128	0,106	0,293
15	25	0,825	0,768	0,099	0,095	0,275
	50	0,809	0,793	0,090	0,080	0,277
	8	0,943	0,903	0,113	0,060	0,504
	10	0,968	0,899	0,119	0,061	0,593
	12	0,952	0,920	0,126	0,094	0,587
	15	0,962	0,910	0,113	0,085	0,565
20	20	0,957	0,900	0,121	0,106	0,554
	25	0,958	0,896	0,117	0,101	0,540
	50	0,934	0,894	0,108	0,100	0,514
	8	0,980	0,950	0,114	0,052	0,699
	10	0,991	0,957	0,140	0,057	0,769
	12	0,993	0,969	0,140	0,087	0,751
25	15	0,992	0,971	0,123	0,095	0,765
	20	0,995	0,959	0,128	0,094	0,745
	25	0,990	0,958	0,120	0,097	0,736
	50	0,983	0,951	0,114	0,112	0,723
	8	0,998	0,982	0,131	0,049	0,831
	10	0,998	0,991	0,135	0,047	0,860
50	12	0,999	0,989	0,133	0,087	0,874
	15	0,998	0,989	0,150	0,115	0,867
	20	1,000	0,997	0,141	0,104	0,861
	25	0,997	0,979	0,123	0,100	0,870
	50	0,996	0,984	0,109	0,101	0,846
	8	1,000	1,000	0,135	0,042	0,993
100	10	1,000	1,000	0,194	0,049	0,996
	12	1,000	1,000	0,217	0,077	0,995
	15	1,000	1,000	0,208	0,104	0,996
	20	1,000	0,999	0,194	0,110	0,992
	25	1,000	1,000	0,152	0,106	0,991
	50	1,000	1,000	0,104	0,106	0,993
150	8	1,000	1,000	0,182	0,080	1,000
	10	1,000	1,000	0,265	0,053	1,000
	12	1,000	1,000	0,299	0,065	1,000
	15	1,000	1,000	0,309	0,110	1,000
	20	1,000	1,000	0,266	0,109	1,000
	25	1,000	1,000	0,215	0,121	1,000
150	50	1,000	1,000	0,095	0,095	1,000
	8	1,000	1,000	0,215	0,119	1,000
	10	1,000	1,000	0,332	0,053	1,000
	12	1,000	1,000	0,402	0,072	1,000
	15	1,000	1,000	0,416	0,100	1,000
	20	1,000	1,000	0,356	0,104	1,000
150	25	1,000	1,000	0,263	0,109	1,000
	50	1,000	1,000	0,136	0,089	1,000

Meta analizinde %40 yanlılık oluşturulduğu,  $OR=2,34$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,70$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 20’de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayırımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, bu teste çok yakın bir performansla Begg testi ve daha sonra Thompson testi takip etmektedir. Harbord ve Schwarzer testleri her koşulda %80 gücün üzerine çıkmadıklarından yanlılığı tespit etme gücü bakımından performanslarının kötü olduğu söylenebilir. Az sayıda çalışma içeren meta analizler ( $k=8$ ) için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür, %80’e ulaşamamıştır. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı meta analizine alınan çalışma sayısı arttıkça yükselmiştir. Begg testi  $k \geq 15$ , Egger testi  $k \geq 15$ , Thompson testi ise  $k \geq 25$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermişlerdir. Schwarzer testinin gücü meta analizine alınan çalışma sayısı fazla olduğunda, meta analizine alınan çalışma sayısının artışından etkilenecek az da olsa bir artış gösterse de %80 güce ulaşamamıştır. Meta analizine alınan çalışma sayısı arttıkça, özellikle çalışma sayısı fazla ve örnekleme alınan birim sayısı az olduğu durumda Harbord testinin gücü yükselmiş fakat %80’e ulaşamamıştır. Çalışmaların örneklem hacminin değişimi ile Begg testinin yanlılığı tespit etme gücü bakımından performansı  $n=10$  değerinden sonra bir miktar düşmüş ve daha sonra sabit kalmıştır. Egger testinin yanlılığı tespit etme gücü bakımından performansı  $n=25$  değerine kadar sabit kalmış, sonra bir miktar düşmüştür. Harbord, Thompson ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansı değişmemiştir. Genel olarak OR değerinin artışıyla birlikte Egger testinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Begg testinin gücü düşmüş, Thompson ve Schwarzer testlerinin performansında bir değişim olmamıştır. OR değerinin artmasıyla Harbord testinin gücünde ise  $k=50,100,150$  olduğu durumlarda bir yükseliş meydana gelmiş fakat güç %80’in üzerine çıkamamıştır.

**Tablo 20.** Meta analizinde %40 yanlışlık oluşturulduğu durumda  $OR=2,34$   $P(S^+|H^+)=0,50$  ve  $P(S^+|H^-)=0,70$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,689	0,707	0,099	0,053	0,114
	10	0,706	0,723	0,092	0,068	0,143
	12	0,707	0,734	0,109	0,077	0,136
	15	0,664	0,715	0,107	0,093	0,163
	20	0,653	0,710	0,108	0,103	0,161
	25	0,633	0,706	0,105	0,103	0,152
10	50	0,608	0,692	0,104	0,097	0,179
	8	0,806	0,815	0,103	0,065	0,227
	10	0,835	0,828	0,098	0,058	0,251
	12	0,798	0,814	0,076	0,065	0,247
	15	0,794	0,816	0,095	0,086	0,273
	20	0,791	0,822	0,101	0,079	0,297
15	25	0,770	0,808	0,105	0,117	0,255
	50	0,718	0,798	0,109	0,097	0,277
	8	0,914	0,923	0,101	0,067	0,444
	10	0,953	0,946	0,076	0,059	0,552
	12	0,929	0,943	0,096	0,058	0,547
	15	0,945	0,948	0,107	0,084	0,566
20	20	0,929	0,929	0,108	0,088	0,574
	25	0,935	0,925	0,104	0,099	0,558
	50	0,899	0,922	0,105	0,096	0,545
	8	0,967	0,967	0,128	0,086	0,655
	10	0,989	0,979	0,075	0,056	0,755
	12	0,979	0,978	0,097	0,070	0,751
25	15	0,979	0,984	0,113	0,079	0,785
	20	0,980	0,970	0,124	0,099	0,771
	25	0,985	0,981	0,113	0,091	0,748
	50	0,960	0,975	0,112	0,104	0,732
	8	0,989	0,990	0,120	0,103	0,811
	10	0,999	0,997	0,090	0,072	0,858
50	12	0,999	0,994	0,097	0,060	0,871
	15	0,995	0,995	0,098	0,060	0,891
	20	0,994	0,989	0,125	0,098	0,866
	25	0,991	0,991	0,105	0,092	0,877
	50	0,990	0,995	0,116	0,104	0,850
	8	1,000	1,000	0,136	0,184	0,997
100	10	1,000	1,000	0,108	0,130	0,999
	12	1,000	1,000	0,114	0,101	0,999
	15	1,000	1,000	0,117	0,072	1,000
	20	1,000	1,000	0,137	0,114	0,999
	25	1,000	1,000	0,139	0,108	0,997
	50	1,000	1,000	0,100	0,093	0,996
150	8	1,000	1,000	0,134	0,341	1,000
	10	1,000	1,000	0,158	0,242	1,000
	12	1,000	1,000	0,132	0,192	1,000
	15	1,000	1,000	0,144	0,083	1,000
	20	1,000	1,000	0,167	0,089	1,000
	25	1,000	1,000	0,153	0,108	1,000
150	50	1,000	1,000	0,122	0,096	1,000
	8	1,000	1,000	0,153	0,510	1,000
	10	1,000	1,000	0,180	0,337	1,000
	12	1,000	1,000	0,132	0,268	1,000
	15	1,000	1,000	0,194	0,085	1,000
	20	1,000	1,000	0,189	0,095	1,000
150	25	1,000	1,000	0,169	0,098	1,000
	50	1,000	1,000	0,130	0,101	1,000

Meta analizinde %40 yanlılık oluşturulduğu,  $OR=4,02$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,80$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 21’de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, bu teste çok yakın bir performansla Begg testi, Thompson testi ve daha sonra sırası ile Harbord ve Schwarzer testleri takip etmektedir. Az sayıda çalışma içeren meta analizler ( $k=8$ ) için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür, %80’e ulaşamamıştır. Begg, Egger, Thompson, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısı arttıkça yükselmiştir. Begg testi  $k \geq 20$ , Egger testi  $k \geq 15$ , Thompson testi ise  $k \geq 50$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermişlerdir. Harbord ve Schwarzer testlerinde örneklemdeki birim sayısının azalması ile güç yükselme eğilimindedir. Meta analizine alınan çalışma sayısı  $k \geq 100$  gibi büyük ve örnekleme alınan birim sayısı az ise Harbord ve Schwarzer testleri güç bakımından iyi performans göstermiştir. Örneklemdeki birim sayısının artması ile Begg, Egger ve Thompson testlerinin güçleri yükselmiştir. Begg, Egger ve Thompson testlerinde genel olarak OR değerinin artışıyla birlikte yanlılığı tespit etme gücü bakımından performans düşmüş, Schwarzer ve Harbord testlerinin performansı yükselmiştir.

**Tablo 21.** Meta analizinde %40 yanlışlık oluşturulduğu durumda  $OR=4,02$   $P(S^+|H^+)=0,50$  ve  $P(S^+|H^-)=0,80$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_k$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,537	0,603	0,193	0,132	0,050
	10	0,621	0,690	0,118	0,108	0,076
	12	0,616	0,696	0,104	0,082	0,084
	15	0,597	0,728	0,090	0,078	0,106
	20	0,640	0,765	0,092	0,077	0,146
	25	0,675	0,769	0,091	0,073	0,145
	50	0,627	0,736	0,108	0,110	0,169
10	8	0,662	0,733	0,252	0,202	0,082
	10	0,726	0,775	0,131	0,136	0,137
	12	0,722	0,798	0,102	0,087	0,151
	15	0,718	0,836	0,097	0,068	0,191
	20	0,794	0,868	0,109	0,069	0,259
	25	0,777	0,863	0,108	0,080	0,283
	50	0,731	0,834	0,108	0,097	0,295
15	8	0,784	0,874	0,329	0,315	0,264
	10	0,890	0,935	0,140	0,189	0,345
	12	0,880	0,924	0,105	0,105	0,364
	15	0,888	0,954	0,079	0,061	0,479
	20	0,932	0,956	0,123	0,081	0,599
	25	0,927	0,970	0,115	0,073	0,593
	50	0,899	0,946	0,096	0,094	0,582
20	8	0,844	0,929	0,444	0,475	0,408
	10	0,952	0,979	0,171	0,246	0,562
	12	0,957	0,980	0,098	0,109	0,598
	15	0,954	0,982	0,077	0,066	0,724
	20	0,981	0,994	0,104	0,070	0,792
	25	0,971	0,989	0,109	0,077	0,804
	50	0,953	0,982	0,099	0,098	0,767
25	8	0,907	0,978	0,500	0,574	0,544
	10	0,985	0,996	0,179	0,298	0,731
	12	0,983	0,996	0,112	0,160	0,757
	15	0,984	0,998	0,083	0,078	0,856
	20	0,992	1,000	0,116	0,075	0,903
	25	0,995	0,992	0,115	0,080	0,916
	50	0,985	0,997	0,108	0,099	0,892
50	8	0,991	1,000	0,756	0,905	0,937
	10	1,000	1,000	0,265	0,554	0,992
	12	1,000	1,000	0,149	0,313	0,991
	15	1,000	1,000	0,110	0,125	1,000
	20	1,000	1,000	0,101	0,076	0,999
	25	1,000	1,000	0,144	0,078	0,999
	50	1,000	1,000	0,106	0,096	1,000
100	8	1,000	1,000	0,937	0,993	0,999
	10	1,000	1,000	0,398	0,835	1,000
	12	1,000	1,000	0,213	0,593	1,000
	15	1,000	1,000	0,103	0,240	1,000
	20	1,000	1,000	0,135	0,117	1,000
	25	1,000	1,000	0,143	0,078	1,000
	50	1,000	1,000	0,131	0,086	1,000
150	8	1,000	1,000	0,982	0,998	1,000
	10	1,000	1,000	0,496	0,952	1,000
	12	1,000	1,000	0,246	0,732	1,000
	15	1,000	1,000	0,120	0,330	1,000
	20	1,000	1,000	0,160	0,140	1,000
	25	1,000	1,000	0,166	0,091	1,000
	50	1,000	1,000	0,144	0,091	1,000

Meta analizinde %40 yanlılık oluşturulduğu,  $OR=9,01$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,90$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 22’de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, Begg, Harbord, Thompson ve Schwarzer testleri takip etmektedir. Bu tez çalışmasında incelenen bütün testlerin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısının artmasıyla yükselmiştir. Begg testi  $k \geq 50$ , Egger testi  $k \geq 20$ , Thompson testi ise  $k \geq 100$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermişlerdir. Thompson testinin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısı ve örneklem hacmi az iken oldukça başarısız olmasına rağmen örnekleme alınan birim sayısı ve çalışma sayısının artması ile oldukça yükselmiştir. Harbord ve Schwarzer testlerinin gücü örnekleme alınan birim sayısı arttıkça düşmüştür. Begg, Egger ve Thompson testlerinin gücü genel olarak örnekleme alınan birim sayısındaki artışa göre yükselme eğiliminde olsa da  $n=12$  olduğu durumda Begg ve Thompson testlerinin performansında bir düşüş görülmüştür. Genel olarak OR değerinin artışıyla birlikte Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları düşmüş, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.



**Tablo 22.** Meta analizinde %40 yanlışlık oluşturulduğu durumda OR=9,01  $P(S^+|H^+)=0,50$  ve  $P(S^+|H^+)=0,90$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,498	0,562	0,537	0,513	0,007
	10	0,534	0,542	0,364	0,431	0,015
	12	0,469	0,494	0,321	0,397	0,010
	15	0,518	0,587	0,168	0,250	0,021
	20	0,594	0,716	0,098	0,171	0,036
	25	0,693	0,826	0,071	0,100	0,060
	50	0,763	0,868	0,105	0,114	0,232
10	8	0,553	0,628	0,653	0,635	0,020
	10	0,596	0,623	0,443	0,552	0,026
	12	0,539	0,538	0,385	0,502	0,023
	15	0,588	0,686	0,219	0,323	0,057
	20	0,716	0,821	0,119	0,223	0,076
	25	0,804	0,905	0,078	0,130	0,141
	50	0,852	0,922	0,109	0,117	0,372
15	8	0,663	0,797	0,796	0,816	0,092
	10	0,725	0,799	0,583	0,742	0,087
	12	0,638	0,705	0,512	0,698	0,072
	15	0,717	0,851	0,269	0,485	0,122
	20	0,881	0,949	0,137	0,301	0,263
	25	0,942	0,981	0,092	0,201	0,445
	50	0,981	0,994	0,113	0,126	0,734
20	8	0,758	0,883	0,897	0,911	0,181
	10	0,804	0,885	0,702	0,849	0,178
	12	0,738	0,836	0,627	0,798	0,136
	15	0,832	0,937	0,313	0,597	0,236
	20	0,949	0,987	0,163	0,458	0,459
	25	0,985	0,999	0,094	0,267	0,680
	50	0,997	0,977	0,111	0,143	0,895
25	8	0,814	0,935	0,940	0,954	0,297
	10	0,872	0,948	0,760	0,915	0,315
	12	0,785	0,897	0,709	0,883	0,226
	15	0,889	0,976	0,384	0,716	0,382
	20	0,978	0,996	0,185	0,544	0,638
	25	0,999	1,000	0,090	0,340	0,842
	50	1,000	1,000	0,101	0,147	0,965
50	8	0,963	0,996	0,997	1,000	0,772
	10	0,975	0,997	0,974	0,998	0,804
	12	0,947	0,990	0,925	0,992	0,663
	15	0,992	1,000	0,626	0,956	0,900
	20	1,000	1,000	0,305	0,815	0,987
	25	1,000	1,000	0,107	0,614	0,999
	50	1,000	1,000	0,119	0,249	1,000
100	8	0,999	1,000	1,000	1,000	0,993
	10	1,000	1,000	1,000	1,000	0,993
	12	0,998	1,000	0,997	1,000	0,997
	15	1,000	1,000	0,887	0,999	0,998
	20	1,000	1,000	0,534	0,980	1,000
	25	1,000	1,000	0,141	0,900	1,000
	50	1,000	1,000	0,159	0,370	1,000
150	8	1,000	1,000	1,000	1,000	1,000
	10	1,000	1,000	1,000	1,000	1,000
	12	1,000	1,000	1,000	1,000	0,998
	15	1,000	1,000	0,967	1,000	1,000
	20	1,000	1,000	0,672	0,998	1,000
	25	1,000	1,000	0,146	0,986	1,000
	50	1,000	1,000	0,214	0,505	1,000

Simülasyon sonucunda (Tablo 18–22) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=8$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Begg, Schwarzer, Harbord ve Thompson testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger testi ise sadece  $OR=9,01$  ve  $n \geq 25$  olduğu durumlarda %80 güce ulaşabilmiştir.

Simülasyon sonucunda (Tablo 18–22) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=10$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi takip etmektedir. Schwarzer, Harbord ve Thompson testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger testi  $OR \geq 2,34$  ve örneklem hacmi büyük olduğu durumlarda %80 güce ulaşabilmiştir. Genel olarak meta analizine alınan çalışma sayısı  $k=8$ 'den  $k=10$ 'a çıktığında Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir, Schwarzer, Harbord ve Thompson testlerinin ise değişmemiştir.

Simülasyon sonucunda (Tablo 18–22) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=15$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra çok da iyi olmayan bir performansla Harbord testi takip etmektedir. Schwarzer ve Thompson testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger ve Begg testi  $OR=1,00$  olduğu durum hariç %80 güce ulaşmış, Harbord testi ise sadece  $OR=9,01$  ve  $n=8$  olduğu durumda %80 güce ulaşabilmiştir. Begg testinin yanlılığı tespit etme gücü performansı  $OR=4,02$  değerinden itibaren özellikle küçük örneklem hacimlerinde düşmeye başlamıştır.

OR=1,50 ve 2,34 iken Begg testi, Egger testi ile aynı performansı göstermiş ve her koşulda %80 güce ulaşmış,  $OR \geq 4,02$  olduğu durumlarda Egger testinden daha kötü performans göstermiştir. OR=1,50, OR=2,34 ve OR=4,02 iken Egger testi her koşulda %80 güce ulaşmıştır. Egger testinin yanlılığı tespit etme gücü performansı OR=9,01 değerinde özellikle küçük örneklem hacimlerinde düşmeye başlamıştır. Genel olarak meta analizine alınan çalışma sayısı k=10'dan k=15'e çıktığında Begg, Egger ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 18–22) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı k=20 olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri sadece OR=9,01 ve örneklem hacmi küçük olduğu durumda %80 güce ulaşabilmiştir. Begg testinin yanlılığı tespit etme gücü  $OR \leq 4,02$  olduğu her durumda %80'e ulaşmış, OR=9,01 olduğu durumda performansı özellikle küçük örneklem hacimlerinde düşmüştür. Egger testinin yanlılığı tespit etme gücü bakımından performansı OR=1,00 hariç her durumda %80'e ulaşmıştır. Meta analizine alınan çalışma sayısı k=15'ten k=20'ye çıktığında bu tez çalışmasında incelenen bütün testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 18–22) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı k=25 olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, Egger testi her koşulda tam olarak %80 güce ulaşarak en iyi performansı göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri sadece OR=9,01 ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Thompson testinin yanlılığı tespit etme gücü bakımından performansı OR=2,34 değerine kadar yükselmiş, daha sonra OR değerinin artışı ile birlikte performansı küçük örneklem hacimlerinde düşmeye başlamıştır. Begg testinin

yanlılığı tespit etme gücü  $OR \leq 4,02$  olduğu her durumda %80'e ulaşmıştır. Genel olarak, meta analizine alınan çalışma sayısı  $k=20$ 'den  $k=25$ 'e çıktığında bu Begg, Egger, Harbord ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Schwarzer testinin performansı değişmemiştir.

Simülasyon sonucunda (Tablo 18–22) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=50$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Begg ve Egger testleri her koşulda tam olarak %80 güce ulaşarak en iyi performansı göstermiştir. Begg ve Egger testlerini, çok yakın bir performansla Thompson testi ve daha sonra Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri sadece  $OR=9,01$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Thompson testinin yanlılığı tespit etme gücü  $OR=9,01$  ve  $n=8$  ve  $12$  olduğu durum hariç her durumda %80'e ulaşmıştır. Genel olarak meta analizine alınan çalışma sayısı  $k=25$ 'ten  $k=50$ 'ye çıktığında sadece Egger testinin yanlılığı tespit etme gücü bakımından performansı değişmemiş, bu tez çalışmasında incelenen diğer testlerin ise yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 18–22) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=100$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Begg, Egger ve Thompson testleri her koşulda tam olarak %80 güce ulaşarak en iyi performansı göstermiştir. Begg, Egger ve Thompson testlerini, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri  $OR \geq 4,02$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Meta analizine alınan çalışma sayısı  $k=50$ 'den  $k=100$ 'e çıktığında bu tez çalışmasında incelenen Thompson, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Begg ve Egger testlerinin ise değişmemiştir.

Simülasyon sonucunda (Tablo 18–22) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=150$  olduğu durumda genel olarak testler güçlerine göre

değerlendirildiğinde, bu tez çalışmasında incelenen Begg, Egger ve Thompson testleri her koşulda tam olarak %80 güce ulaşarak en iyi performansı göstermiştir. Begg, Egger ve Thompson testlerini, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri  $OR \geq 4,02$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Genel olarak meta analizine alınan çalışma sayısı  $k=100$ 'den  $k=150$ 'ye çıktığında bu tez çalışmasında incelenen bütün testlerin yanlılığı tespit etme gücü bakımından performansları değişmemiştir.

Simülasyon sonucunda (Tablo 18–22) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve örneklem hacmi  $n=8$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=4,02$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren düşmeye başlamıştır. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=4,02$  değerinden itibaren yükselmiştir.

Simülasyon sonucunda (Tablo 18–22) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve örneklem hacmi  $n=10$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Begg ve Egger testleri göstermiştir Begg ve Egger testlerini, Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=4,02$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren düşmeye başlamıştır. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=4,02$  değerinden itibaren yükselmiştir. Genel olarak örnekleme alınan birim sayısının  $n=8$ 'den,  $n=10$ 'a çıkmasıyla Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselirken, Harbord ve Schwarzer testlerinin azalmıştır.

Simülasyon sonucunda (Tablo 18–22) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve örneklem hacmi  $n=12$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg testinin yanlılığı tespit etme gücü bakımından performansı  $OR=1,50$  değerine kadar yükseliş eğiliminde iken bu değerden sonra düşmeye başlamıştır. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=4,02$  değerinden sonra yükselmiştir. Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=4,02$  değerine kadar yükseliş eğiliminde iken, bu değerden itibaren düşmeye başlamıştır. Genel olarak örnekleme alınan birim sayısının  $n=10$ 'dan,  $n=12$ 'ye çıkmasıyla Begg, Egger, Harbord ve Thompson testlerinin testinin yanlılığı tespit etme gücü bakımından performansları düşmüş, Schwarzer testinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 18–22) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve örneklem hacmi  $n=15$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg testinin yanlılığı tespit etme gücü bakımından performansı  $OR=1,50$  değerine kadar yükseliş eğiliminde iken bu değerden sonra düşme eğilimi göstermiştir. Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı  $OR=9,01$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren düşme eğilimi göstermiştir. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=9,01$  değerinden itibaren yükselmiştir. Genel olarak örnekleme alınan birim sayısının  $n=12$ 'den,  $n=15$ 'e çıkmasıyla Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Schwarzer ve Harbord testlerinin performansları düşmüştür.

Simülasyon sonucunda (Tablo 18–22) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve örneklem hacmi  $n=20$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson ve Harbord testi takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren test Schwarzer testidir. Schwarzer testi hiçbir durumda %80 güce ulaşamamıştır. Begg testinin yanlılığı tespit etme gücü bakımından performansı  $OR=1,50$  değerine kadar yükseliş eğiliminde iken bu değerden sonra düşme eğilimi göstermiş ve daha sonra değişmemiştir. Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=9,01$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren Thompson testinin performansı düşmeye başlamıştır. Harbord testinin yanlılığı tespit etme gücü bakımından performansı ise  $OR=9,01$  değerinden itibaren yükselmiştir. Genel olarak örnekleme alınan birim sayısının  $n=15$ 'ten,  $n=20$ 'ye çıkmasıyla Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Schwarzer testinin performansı düşmüş, Harbord ve Thompson testlerinin performansları değişmemiştir.

Simülasyon sonucunda (Tablo 18–22) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlılık oluşturulduğu ve örneklem hacmi  $n=25$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson ve Harbord testi takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren test Schwarzer testidir. Schwarzer testi hiçbir durumda %80 güce ulaşamamıştır. OR değeri arttıkça Begg, Egger, Thompson ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğilimindedir. Egger testi  $OR=9,01$  olduğu her durumda %80 güce ulaşmıştır. Genel olarak örnekleme alınan birim sayısının  $n=20$ 'den,  $n=25$ 'e çıkmasıyla Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Harbord testinin performansı düşmüş, Schwarzer testinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 18–22) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %40 yanlışlık oluşturulduğu ve örneklem hacmi  $n=50$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Begg ve Egger testleri göstermiştir. Begg ve Egger testlerini, Thompson testi takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren, hiçbir durumda %80 güce ulaşamayan testler Harbord ve Schwarzer testleridir. OR değeri arttıkça Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğilimindedir. Egger testi  $OR=9,01$  olduğu her durumda %80 güce ulaşmıştır. Genel olarak örnekleme alınan birim sayısının  $n=25$ 'ten,  $n=50$ 'ye çıkmasıyla Egger ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları düşmüş, Thompson, Begg ve Schwarzer testlerinin performansı ise değişmemiştir.

Meta analizinde %40 yanlışlık oluşturulduğu durum dikkate alınarak yapılan simülasyon senaryoları (Tablo 18–22) incelendiğinde örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayrımı yapılmaksızın genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. OR değerinin artmasıyla Begg testinin yanlılığı tespit etme gücü bakımından performansı  $OR= 2,34$  değerine kadar yükselmiş, bu değerden itibaren düşmüştür. Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=4,02$  değerine kadar yükselmiş, bu OR değerinden itibaren düşmeye başlamıştır. Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları, OR değerinin artmasıyla özellikle  $OR=4,02$  değerinden sonra, örneklem hacminin azalması ve meta analizine alınan çalışma sayılarının artması ile yükselmiştir. Bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları meta analizine alınan çalışma sayıları arttıkça yükselmiştir. Meta analizine alınan çalışma sayısı az iken bu tez çalışmasında kullanılan testlerin hiçbiri yanlılığı tespit etme gücü bakımından iyi performans göstermezken, meta analizine alınan çalışma sayısı yüksek ise Begg, Egger ve



Thompson testleri çok iyi performans göstermiştir. Meta analizine alınan çalışma sayısı  $k=20,25,50$  olduğu durumlarda da Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları iyidir. Örneklem hacmi arttıkça Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselme eğiliminde iken, Schwarzer ve Harbord testlerinin performansları düşme eğilimindedir.

#### **4.5. Meta Analizinde %60 Yanlılık Oluşturduğu Duruma Ait Bulgular**

Meta analizine alınan çalışmaların OR değerleri sıralanıp en büyük OR değerlerine sahip olanların %60'ı meta analizinden çıkarılarak funnel grafiğinin sağ tarafında asimetri sağlanması ile yanlılık oluşturulduğu durum dikkate alınarak yapılan simülasyon çalışması ile elde edilen veriler analiz edilmiş ve analizlerden elde edilen bulgular tablolar (Tablo 23–27) aracılığıyla sunulmuştur. Hasta grubunda pozitif sonuç oranlarına göre ayrı ayrı sunulan tablolarda anakütle OR değerleri,  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  olan örneklem hacimleri ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  olan meta analizine alınan çalışma sayılarına göre funnel grafiği asimetrisinin değerlendirilmesi ile meta analizinde yanlılık olup olmadığına karar vermemizi sağlayan Begg, Egger, Schwarzer, Harbord ve Thompson testlerine ait yanlılığı tespit etme güçleri verilmiştir. Nominal değer yani Tip-I hata düzeyi  $\alpha=0,10$  olarak alınmıştır.

Meta analizinde %60 yanlılık oluşturulduğu,  $OR=1,00$ ,  $P(S^+ \setminus H^-)=0,50$  ve  $P(S^+ \setminus H^+)=0,50$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 23'te gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Begg testi göstermiştir. Begg testini, bu teste çok yakın bir performans ile Egger testi, daha sonra sırası ile Thompson, Schwarzer testleri takip etmektedir. Harbord testi her koşulda %80 gücün üzerine çıkamadığından yanlılığı tespit etme gücü bakımından performansının kötü olduğu söylenebilir. Az sayıda

çalışma içeren meta analizler için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür, %80'e ulaşamamıştır. Begg, Egger, Thompson ve Schwarzer testlerinin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısı arttıkça yükselmiştir. Begg testi  $k \geq 15$ , Egger testi  $k \geq 20$ , Thompson testi ise  $k \geq 50$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermişlerdir. Schwarzer testinin yanlılığı tespit etme gücü sadece  $k=150$  ve  $n=12, 15, 20$  olduğu durumlarda %80'in üzerine çıkmıştır. Harbord testinin gücü meta analizine alınan çalışma sayısının artışıyla, çalışma sayısı fazla ise az da olsa etkilenerek artış göstermiştir. Begg, Egger, Thompson, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü örneklem hacminin değişiminden etkilenmemiştir.

**Tablo 23.** Meta analizinde %60 yanlışlık oluşturulduğu durumda OR=1,00  $P(S^+|H^+)=0,50$  ve  $P(S^+|H^+)=0,50$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,652	0,636	0,129	0,085	0,097
	10	0,635	0,590	0,118	0,106	0,107
	12	0,620	0,582	0,128	0,086	0,091
	15	0,628	0,594	0,131	0,090	0,102
	20	0,606	0,562	0,126	0,099	0,094
	25	0,603	0,575	0,101	0,080	0,094
10	50	0,537	0,550	0,121	0,101	0,083
	8	0,739	0,710	0,167	0,091	0,191
	10	0,734	0,678	0,148	0,090	0,203
	12	0,740	0,671	0,138	0,084	0,182
	15	0,743	0,666	0,152	0,098	0,210
	20	0,709	0,654	0,119	0,088	0,187
15	25	0,725	0,672	0,126	0,106	0,188
	50	0,641	0,626	0,134	0,111	0,159
	8	0,906	0,833	0,160	0,072	0,393
	10	0,909	0,839	0,150	0,088	0,407
	12	0,884	0,832	0,156	0,089	0,408
	15	0,904	0,815	0,174	0,107	0,438
20	20	0,878	0,818	0,160	0,119	0,416
	25	0,897	0,803	0,162	0,103	0,368
	50	0,844	0,767	0,152	0,127	0,377
	8	0,966	0,917	0,202	0,068	0,600
	10	0,969	0,925	0,216	0,102	0,611
	12	0,955	0,914	0,192	0,101	0,607
25	15	0,955	0,898	0,222	0,110	0,625
	20	0,954	0,894	0,205	0,118	0,571
	25	0,964	0,884	0,194	0,105	0,563
	50	0,943	0,860	0,160	0,110	0,540
	8	0,989	0,952	0,204	0,073	0,720
	10	0,990	0,959	0,231	0,105	0,747
50	12	0,983	0,939	0,218	0,109	0,747
	15	0,991	0,959	0,236	0,108	0,727
	20	0,985	0,931	0,240	0,119	0,699
	25	0,985	0,936	0,227	0,103	0,686
	50	0,967	0,905	0,182	0,106	0,644
	8	1,000	1,000	0,309	0,070	0,969
100	10	0,999	1,000	0,388	0,118	0,972
	12	1,000	0,996	0,412	0,140	0,968
	15	1,000	0,995	0,427	0,142	0,971
	20	1,000	0,999	0,382	0,157	0,959
	25	1,000	0,999	0,383	0,130	0,958
	50	0,999	0,993	0,283	0,113	0,935
150	8	1,000	1,000	0,475	0,059	1,000
	10	1,000	1,000	0,598	0,151	1,000
	12	1,000	1,000	0,692	0,211	1,000
	15	1,000	1,000	0,732	0,206	1,000
	20	1,000	1,000	0,648	0,176	0,999
	25	1,000	1,000	0,632	0,157	1,000
150	50	1,000	1,000	0,451	0,126	1,000
	8	1,000	1,000	0,586	0,070	1,000
	10	1,000	1,000	0,737	0,183	1,000
	12	1,000	1,000	0,829	0,248	1,000
	15	1,000	1,000	0,860	0,264	1,000
	20	1,000	1,000	0,806	0,246	1,000
150	25	1,000	1,000	0,757	0,201	1,000
	50	1,000	1,000	0,599	0,145	1,000

Meta analizinde %60 yanlılık oluşturulduğu,  $OR=1,50$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,60$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 24’te gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayırımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Begg testi göstermiştir. Begg testini, bu teste çok yakın bir performansla Egger testi ve daha sonra Thompson testi takip etmektedir. Harbord ve Schwarzer testleri her koşulda %80 gücün üzerine çıkamadıklarından yanlılığı tespit etme gücü bakımından performanslarının kötü olduğu söylenebilir. Az sayıda çalışma içeren meta analizler ( $k=8$ ) için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür, %80’e ulaşamamıştır. Begg, Egger, ve Thompson testlerinin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısı arttıkça yükselmiş, Harbord ve Schwarzer testlerinin gücü ise meta analizine alınan çalışma sayısının artışıyla etkilenerek yükseliş gösterse de %80 güce ulaşamamıştır. Begg testi  $k \geq 10$ , Egger testi  $k \geq 15$ , Thompson testi ise  $k \geq 50$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermiştir. Bu tez çalışmasında ele alınan Begg, Schwarzer ve Harbord testlerinin gücünün örneklem hacmine göre değişmediği söylenebilir. Egger ve Thompson testlerinin gücü az da olsa azalış eğilimindedir. Genel olarak OR değerinin artışıyla birlikte Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü yükselmiş, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü düşmüştür.

**Tablo 24.** Meta analizinde %60 yanlışlık oluşturulduğu durumda OR=1,50  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,60$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,721	0,711	0,117	0,064	0,088
	10	0,755	0,735	0,109	0,082	0,140
	12	0,733	0,719	0,128	0,101	0,112
	15	0,724	0,710	0,098	0,103	0,142
	20	0,723	0,686	0,093	0,095	0,122
	25	0,688	0,665	0,101	0,105	0,105
10	8	0,687	0,677	0,114	0,106	0,112
	10	0,846	0,799	0,120	0,072	0,177
	12	0,879	0,829	0,104	0,066	0,248
	15	0,861	0,809	0,108	0,086	0,234
	20	0,846	0,794	0,101	0,096	0,245
	25	0,833	0,775	0,106	0,092	0,252
15	8	0,810	0,759	0,102	0,097	0,226
	10	0,803	0,781	0,101	0,100	0,215
	12	0,957	0,926	0,122	0,059	0,439
	15	0,976	0,939	0,115	0,063	0,539
	20	0,948	0,922	0,130	0,087	0,520
	25	0,962	0,922	0,129	0,100	0,511
20	8	0,952	0,897	0,111	0,109	0,483
	10	0,966	0,908	0,109	0,109	0,486
	12	0,935	0,904	0,107	0,103	0,469
	15	0,987	0,962	0,112	0,053	0,659
	20	0,986	0,964	0,126	0,060	0,725
	25	0,990	0,971	0,114	0,087	0,727
25	8	0,990	0,970	0,124	0,106	0,715
	10	0,987	0,957	0,113	0,102	0,684
	12	0,986	0,964	0,113	0,097	0,662
	15	0,982	0,950	0,102	0,088	0,663
	20	0,999	0,989	0,097	0,058	0,818
	25	0,999	0,969	0,116	0,059	0,850
50	8	0,998	0,991	0,124	0,080	0,826
	10	0,997	0,989	0,133	0,102	0,838
	12	0,999	0,981	0,117	0,102	0,801
	15	0,999	0,969	0,122	0,100	0,784
	20	0,994	0,982	0,112	0,102	0,775
	25	1,000	1,000	0,096	0,039	0,995
100	8	1,000	1,000	0,160	0,049	0,997
	10	1,000	1,000	0,173	0,074	0,997
	12	1,000	1,000	0,141	0,100	0,994
	15	1,000	0,999	0,156	0,129	0,991
	20	1,000	1,000	0,127	0,116	0,986
	25	1,000	1,000	0,119	0,090	0,992
150	8	1,000	1,000	0,094	0,077	1,000
	10	1,000	1,000	0,248	0,045	1,000
	12	1,000	1,000	0,267	0,069	1,000
	15	1,000	1,000	0,231	0,112	1,000
	20	1,000	1,000	0,164	0,134	1,000
	25	1,000	1,000	0,153	0,124	1,000
150	8	1,000	1,000	0,123	0,103	1,000
	10	1,000	1,000	0,125	0,098	1,000
	12	1,000	1,000	0,333	0,043	1,000
	15	1,000	1,000	0,338	0,067	1,000
	20	1,000	1,000	0,251	0,115	1,000
	25	1,000	1,000	0,156	0,151	1,000
150	8	1,000	1,000	0,211	0,133	1,000
	10	1,000	1,000	0,211	0,133	1,000
	12	1,000	1,000	0,211	0,133	1,000
	15	1,000	1,000	0,211	0,133	1,000
	20	1,000	1,000	0,211	0,133	1,000
	25	1,000	1,000	0,162	0,115	1,000

Meta analizinde %60 yanlılık oluşturulduğu,  $OR=2,34$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,70$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 25'te gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, bu teste çok yakın bir performansla Begg testi ve daha sonra Thompson ve Harbord testi takip etmektedir. Thompson testi  $k \geq 25$  olduğu durumlarda %80'nin üzerinde güç değerleri elde etmiştir. Schwarzer testi her koşulda %80 gücün üzerine çıkamadığından yanlılığı tespit etme gücü bakımından performansının kötü olduğu söylenebilir. Az sayıda çalışma içeren meta analizler ( $k=8$ ) için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür, %80'e ulaşamamıştır. Begg, Egger, Thompson ve Harbord testlerinin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısı arttıkça yükselmiştir. Begg testi  $k \geq 15$ , Egger testi  $k \geq 15$ , Thompson testi ise  $k \geq 50$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermiştir. Çalışmaların örneklem hacminin değişimi ile Begg testinin yanlılığı tespit etme gücü bakımından performansı  $n=12$  değerinden sonra bir miktar düşmüş ve daha sonra sabit kalmıştır. Egger testinin yanlılığı tespit etme gücü bakımından performansı  $n=25$  değerine kadar sabit kalmış, sonra bir miktar düşmüştür. Thompson testinin yanlılığı tespit etme gücü bakımından performansı örneklem hacminin artışı ile önce yükselmiş sonra düşmüştür. Çalışmaların örneklem hacminin değişimi ile Schwarzer ve Harbord testlerinin gücü çalışma sayısı arttıkça ve örneklem hacmi  $n=8$  olduğu durumlarda belirgin bir şekilde yükselmiş, fakat Schwarzer testinin gücü %80'e ulaşamamıştır. Genel olarak OR değerinin artışıyla birlikte Egger ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Begg ve Thompson testlerinin performansı düşmüştür. OR değerinin artışıyla birlikte Schwarzer testinin gücü meta analizine alınan çalışma sayısı fazla ve örneklem hacmi az ise yükselmiş, diğer durumlarda gözle görülür bir değişim meydana gelmemiştir.

**Tablo 25.** Meta analizinde %60 yanlışlık oluşturulduğu durumda  $OR=2,34$   $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,70$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,694	0,718	0,107	0,053	0,057
	10	0,702	0,713	0,093	0,067	0,091
	12	0,675	0,711	0,109	0,081	0,086
	15	0,667	0,721	0,116	0,096	0,112
	20	0,675	0,723	0,118	0,104	0,107
	25	0,626	0,691	0,095	0,093	0,102
	50	0,588	0,658	0,107	0,109	0,115
10	8	0,783	0,800	0,143	0,065	0,152
	10	0,832	0,805	0,079	0,066	0,181
	12	0,803	0,829	0,100	0,074	0,197
	15	0,785	0,826	0,112	0,083	0,214
	20	0,776	0,804	0,121	0,099	0,218
	25	0,770	0,802	0,092	0,094	0,218
	50	0,715	0,789	0,125	0,112	0,215
15	8	0,885	0,915	0,152	0,094	0,344
	10	0,946	0,945	0,084	0,066	0,428
	12	0,931	0,947	0,096	0,069	0,455
	15	0,932	0,939	0,114	0,085	0,487
	20	0,930	0,932	0,103	0,095	0,489
	25	0,942	0,933	0,098	0,100	0,500
	50	0,863	0,908	0,139	0,114	0,471
20	8	0,946	0,964	0,187	0,147	0,545
	10	0,989	0,980	0,085	0,069	0,653
	12	0,980	0,982	0,084	0,056	0,672
	15	0,975	0,978	0,117	0,080	0,701
	20	0,985	0,979	0,108	0,082	0,689
	25	0,981	0,979	0,115	0,085	0,695
	50	0,959	0,963	0,111	0,092	0,646
25	8	0,976	0,990	0,207	0,211	0,687
	10	0,994	0,996	0,101	0,073	0,793
	12	0,994	0,995	0,098	0,070	0,841
	15	0,996	0,995	0,109	0,079	0,846
	20	0,994	0,994	0,122	0,094	0,844
	25	0,995	0,993	0,127	0,099	0,826
	50	0,980	0,992	0,114	0,098	0,787
50	8	0,999	1,000	0,341	0,502	0,980
	10	1,000	1,000	0,116	0,160	0,989
	12	1,000	1,000	0,110	0,071	0,993
	15	1,000	1,000	0,119	0,072	0,996
	20	1,000	0,999	0,135	0,099	0,995
	25	1,000	1,000	0,136	0,091	0,997
	50	1,000	1,000	0,118	0,089	0,992
100	8	1,000	1,000	0,553	0,819	1,000
	10	1,000	1,000	0,188	0,385	1,000
	12	1,000	1,000	0,123	0,101	1,000
	15	1,000	1,000	0,148	0,075	1,000
	20	1,000	1,000	0,161	0,085	1,000
	25	1,000	1,000	0,155	0,093	1,000
	50	1,000	1,000	0,145	0,097	1,000
150	8	1,000	1,000	0,725	0,936	1,000
	10	1,000	1,000	0,216	0,593	1,000
	12	1,000	1,000	0,159	0,116	1,000
	15	1,000	1,000	0,151	0,108	1,000
	20	1,000	1,000	0,150	0,079	1,000
	25	1,000	1,000	0,166	0,114	1,000
	50	1,000	1,000	0,164	0,103	1,000

Meta analizinde %60 yanlılık oluşturulduğu,  $OR=4,02$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,80$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 26’da gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, bu teste çok yakın bir performansla Begg testi, Thompson testi ve daha sonra sırası ile Harbord ve Schwarzer testleri takip etmektedir. Az sayıda çalışma içeren meta analizler ( $k=8$ ) için bu tez çalışmasında kullanılan bütün testlerin gücü oldukça düşüktür, %80’e ulaşamamıştır. Begg, Egger, Thompson, Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısı arttıkça yükselmiştir. Begg testi  $k \geq 20$ , Egger testi  $k \geq 15$ , Thompson testi ise  $k \geq 50$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermiştir. Örneklemdeki birim sayısının artması ile Begg, Egger ve Thompson testlerinin güçleri yükselme eğilimindedir. Harbord ve Schwarzer testlerinde örneklemdeki birim sayısının azalması ile güç yükselme eğilimindedir. Meta analizine alınan çalışma sayısı  $k \geq 50$  gibi büyük ve örnekleme alınan birim sayısı  $n \leq 10$  gibi küçük değerler aldıkça Harbord testi güç bakımından iyi performans göstermiştir. Meta analizine alınan çalışma sayısı  $k \geq 100$  ve örnekleme alınan birim sayısı  $n=8$  ise Schwarzer testi güç bakımından iyi performans göstermiştir. OR değerinin artışıyla birlikte Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü örneklem hacmi az ise yükselmiştir. Begg, Egger ve Thompson testlerinde OR değerinin artışıyla birlikte genel olarak örneklemdeki birim sayısı az ise güç düşme, örneklemdeki birim sayısı fazla ise güç yükselme eğilimindedir.



**Tablo 26.** Meta analizinde %60 yanlışlık oluşturulduğu durumda OR=4,02  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,80$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,571	0,623	0,232	0,189	0,014
	10	0,632	0,623	0,114	0,119	0,027
	12	0,651	0,681	0,104	0,084	0,037
	15	0,577	0,686	0,096	0,080	0,059
	20	0,645	0,755	0,094	0,069	0,085
	50	0,633	0,733	0,094	0,086	0,092
10	8	0,589	0,707	0,086	0,089	0,112
	10	0,655	0,722	0,262	0,198	0,045
	12	0,747	0,766	0,123	0,139	0,087
	15	0,744	0,783	0,101	0,099	0,086
	20	0,703	0,805	0,079	0,074	0,128
	50	0,781	0,844	0,104	0,071	0,197
15	8	0,717	0,805	0,098	0,091	0,217
	10	0,779	0,872	0,385	0,357	0,171
	12	0,884	0,911	0,142	0,208	0,240
	15	0,865	0,908	0,106	0,122	0,272
	20	0,867	0,940	0,073	0,068	0,362
	50	0,932	0,958	0,111	0,066	0,484
20	8	0,910	0,959	0,122	0,080	0,509
	10	0,856	0,949	0,485	0,515	0,338
	12	0,952	0,968	0,185	0,312	0,420
	15	0,953	0,970	0,104	0,139	0,467
	20	0,944	0,983	0,085	0,083	0,589
	50	0,984	0,987	0,095	0,073	0,692
25	8	0,968	0,989	0,142	0,094	0,726
	10	0,959	0,978	0,104	0,092	0,685
	12	0,911	0,976	0,556	0,587	0,501
	15	0,975	0,992	0,209	0,397	0,591
	20	0,981	0,991	0,115	0,189	0,636
	50	0,981	0,996	0,091	0,101	0,772
50	8	0,994	0,994	0,125	0,083	0,842
	10	0,994	0,994	0,112	0,090	0,824
	12	0,994	1,000	0,735	0,847	0,943
	15	0,999	1,000	0,327	0,739	0,961
	20	1,000	1,000	0,143	0,413	0,975
	50	1,000	1,000	0,083	0,180	0,998
100	8	1,000	1,000	0,102	0,082	1,000
	10	1,000	1,000	0,147	0,094	0,996
	12	1,000	1,000	0,120	0,082	0,999
	15	1,000	1,000	0,879	0,975	1,000
	20	1,000	1,000	0,481	0,973	1,000
	50	1,000	1,000	0,253	0,732	1,000
150	8	1,000	1,000	0,126	0,415	1,000
	10	1,000	1,000	0,129	0,136	1,000
	12	1,000	1,000	0,129	0,136	1,000
	15	1,000	1,000	0,184	0,102	1,000
	20	1,000	1,000	0,150	0,083	1,000
	50	1,000	1,000	0,950	0,993	1,000
150	8	1,000	1,000	0,594	0,998	1,000
	10	1,000	1,000	0,327	0,882	1,000
	12	1,000	1,000	0,176	0,602	1,000
	15	1,000	1,000	0,173	0,162	1,000
	20	1,000	1,000	0,206	0,125	1,000
	50	1,000	1,000	0,174	0,089	1,000

Meta analizinde %60 yanlılık oluşturulduğu,  $OR=9,01$ ,  $P(S^+|H^-)=0,50$  ve  $P(S^+|H^+)=0,90$  olduğu durumda  $n_H=n_K= 8, 10, 12, 15, 20, 25, 50$  ve  $k= 8, 10, 15, 20, 25, 50, 100, 150$  için yapılan karşılaştırmalar neticesinde testlerin güçleri yönünden performanslarına ait simülasyon sonuçları Tablo 27’de gösterilmiştir. Simülasyon sonucunda örneklem hacmi ve meta analizine alınan çalışma sayıları ayrımı yapılmaksızın testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, genel olarak en iyi performansı Egger testi göstermiştir. Egger testini, Begg, Harbord, Thompson ve Schwarzer testleri takip etmektedir. Bu tez çalışmasında incelenen bütün testlerin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısının artmasıyla yükselmiştir. Begg testi  $k \geq 50$ , Egger testi  $k \geq 25$ , Thompson testi ise  $k \geq 100$  olduğu her durumda %80 gücün üzerine çıkarak iyi performans göstermiştir. Thompson testinin yanlılığı tespit etme gücü meta analizine alınan çalışma sayısı ve örneklem hacmi az iken oldukça başarısız olmasına rağmen örnekleme alınan birim sayısı ve çalışma sayısının artması ile yükselmiştir. Harbord ve Schwarzer testlerinin gücü örnekleme alınan birim sayısı arttıkça düşmüştür. Begg ve Egger testlerinin gücünde genel olarak örnekleme alınan birim sayısındaki artışa göre önce düşüş sonra yükseliş olduğu söylenebilir. Thompson testinin yanlılığı tespit etme gücü bakımından performansı örneklem hacminin artmasıyla yükselmiştir. Genel olarak OR değerinin artışıyla birlikte Schwarzer ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiştir. OR değerinin artışıyla birlikte Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansında düşme meydana gelmiştir.

**Tablo 27.** Meta analizinde %60 yanlışlık oluşturulduğu durumda OR=9,01  $P(S^+|H^+)=0,50$  ve  $P(S^+|H^+)=0,90$  için testlerin yanlışlığı tespit etme gücüne ilişkin simülasyon sonuçları

k	$n_H=n_K$	Begg Yöntemi	Egger Yöntemi	Schwarzer Yöntemi	Harbord Yöntemi	Thompson Yöntemi
8	8	0,617	0,636	0,569	0,568	0,002
	10	0,622	0,556	0,401	0,504	0,003
	12	0,489	0,424	0,385	0,490	0,002
	15	0,550	0,515	0,195	0,333	0,007
	20	0,588	0,639	0,110	0,224	0,012
	25	0,686	0,773	0,073	0,133	0,020
	50	0,744	0,843	0,094	0,118	0,136
10	8	0,660	0,696	0,681	0,692	0,006
	10	0,679	0,626	0,495	0,585	0,012
	12	0,580	0,508	0,427	0,569	0,007
	15	0,618	0,622	0,244	0,415	0,016
	20	0,720	0,762	0,113	0,289	0,038
	25	0,813	0,877	0,084	0,161	0,074
	50	0,861	0,921	0,101	0,133	0,287
15	8	0,773	0,838	0,808	0,844	0,055
	10	0,778	0,779	0,630	0,797	0,058
	12	0,673	0,643	0,583	0,781	0,027
	15	0,754	0,806	0,317	0,584	0,045
	20	0,834	0,906	0,160	0,433	0,134
	25	0,943	0,975	0,079	0,244	0,275
	50	0,969	0,994	0,108	0,134	0,639
20	8	0,843	0,910	0,901	0,932	0,138
	10	0,843	0,869	0,737	0,893	0,116
	12	0,756	0,757	0,697	0,891	0,065
	15	0,849	0,901	0,386	0,719	0,119
	20	0,924	0,972	0,179	0,581	0,256
	25	0,987	0,997	0,090	0,327	0,519
	50	0,994	0,997	0,118	0,130	0,850
25	8	0,909	0,958	0,952	0,977	0,244
	10	0,910	0,934	0,832	0,958	0,180
	12	0,797	0,844	0,780	0,944	0,099
	15	0,876	0,962	0,451	0,808	0,188
	20	0,960	0,991	0,207	0,694	0,422
	25	0,998	1,000	0,086	0,398	0,720
	50	0,998	1,000	0,128	0,186	0,942
50	8	0,985	0,999	0,997	0,999	0,698
	10	0,985	1,000	0,975	1,000	0,606
	12	0,932	0,981	0,950	0,996	0,408
	15	0,985	0,999	0,743	0,987	0,707
	20	1,000	1,000	0,330	0,930	0,941
	25	1,000	1,000	0,101	0,732	0,994
	50	0,998	1,000	0,140	0,302	1,000
100	8	0,997	1,000	1,000	1,000	0,987
	10	1,000	1,000	1,000	1,000	0,964
	12	0,984	1,000	0,997	1,000	0,869
	15	1,000	1,000	0,930	1,000	0,992
	20	1,000	1,000	0,588	0,998	1,000
	25	1,000	1,000	0,122	0,951	1,000
	50	0,998	1,000	0,196	0,463	1,000
150	8	1,000	1,000	1,000	1,000	1,000
	10	1,000	1,000	1,000	1,000	0,996
	12	0,998	1,000	1,000	1,000	0,979
	15	1,000	1,000	0,987	1,000	1,000
	20	1,000	1,000	0,733	1,000	1,000
	25	1,000	1,000	0,122	0,997	1,000
	50	0,998	1,000	0,245	0,628	1,000

Simülasyon sonucunda (Tablo 23–27) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=8$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Begg, Schwarzer, Harbord ve Thompson testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Egger testi ise sadece  $OR= 9,01$  ve  $n=50$  olduğu durumda %80 güce ulaşabilmiştir.

Simülasyon sonucunda (Tablo 23–27) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=10$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi takip etmektedir. Schwarzer, Harbord ve Thompson testleri yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir, hiçbiri %80 güce ulaşamamıştır. Eger ve Begg testleri  $OR=1,00$  olduğu durumda %80 güce ulaşamamıştır.  $OR=2,34$  değerine kadar OR değerinin artışı ile Egger testinin yanlılığı tespit etme gücü bakımından performansı özellikle küçük örneklem hacimlerinde yükselmiş,  $OR=2,34$  değerinden sonra büyük örneklem hacimlerinde iyi performans göstererek azalmaya başlamıştır. Begg testi  $OR=1,50$  olduğu her durumda %80 güce ulaşarak iyi performans göstermiş, daha sonra OR değerinin artışı ile önce sadece küçük örneklem hacimlerinde, sonra sadece büyük örneklem hacimlerinde %80 güce ulaşabilmiş ve testin performansı düşmüştür. Genel olarak meta analizine alınan çalışma sayısı  $k=8$ 'den  $k=10$ 'a çıktığında Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Schwarzer, Harbord ve Thompson testlerinin ise değişmemiştir.

Simülasyon sonucunda (Tablo 23–27) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=15$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra çok da iyi olmayan bir

performansla Harbord ve Schwarzer testleri takip etmektedir. Thompson testi yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testtir, %80 güce ulaşamamıştır. Harbord ve Schwarzer testleri sadece  $OR=9,01$  ve  $n=8$  olduğu durumda %80 güce ulaşabilmiştir. Begg testinin yanlılığı tespit etme gücü performansı  $OR=4,02$  değerinden itibaren özellikle küçük örneklem hacimlerinde düşmeye başlamıştır.  $OR \leq 2,34$  iken Begg testi her koşulda %80 güce ulaşmış,  $OR \geq 4,02$  olduğu durumlarda Egger testinden daha kötü performans göstermiştir.  $OR=1,50$ ,  $OR=2,34$  ve  $OR=4,02$  iken Egger testi ile her koşulda %80 güce ulaşmıştır. Egger testinin yanlılığı tespit etme gücü performansı  $OR=9,01$  değerinde özellikle küçük örneklem hacimlerinde düşmeye başlamıştır. Genel olarak meta analizine alınan çalışma sayısı  $k=10$ 'dan  $k=15$ 'e çıktığında Begg, Egger, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 23–27) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=20$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Begg ve Egger testleri göstermiştir. Begg ve Egger testlerini, Harbord, Thompson ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri sadece  $OR=9,01$  ve örneklem hacmi küçük olduğu durumda, Thompson testi ise sadece  $OR=9,01$  ve örneklem hacmi büyük olduğu durumda %80 güce ulaşabilmiştir. Begg ve Egger testlerinin yanlılığı tespit etme gücü  $OR \leq 4,02$  olduğu her durumda %80'e ulaşmıştır. Meta analizine alınan çalışma sayısı  $k=15$ 'ten  $k=20$ 'ye çıktığında Begg, Egger, Harbord ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 23–27) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=25$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, Egger testinin her koşulda tam olarak %80 güce ulaşmasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer ve Harbord testleri sadece  $OR=9,01$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Thompson testinin yanlılığı tespit etme gücü

bakımından performansı  $OR=1,50$  değerine kadar yükselmiş, daha sonra  $OR$  değerinin artışı ile birlikte performansı küçük örneklem hacimlerinde düşmeye başlamıştır. Begg testinin yanlılığı tespit etme gücü  $OR \leq 4,02$  olduğu her durumda %80'e ulaşmıştır. Genel olarak, meta analizine alınan çalışma sayısı  $k=20$ 'den  $k=25$ 'e çıktığında, Egger, Harbord, Schwarzer ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Begg testinin performansı değişmemiştir.

Simülasyon sonucunda (Tablo 23–27) örneklem hacmi ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=50$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Begg ve Egger testlerinin her koşulda tam olarak %80 güce ulaşmasıyla birlikte, en iyi performansı Begg ve Egger göstermiştir. Begg ve Egger testlerini, çok yakın bir performansla Thompson testi ve daha sonra Harbord ve Schwarzer testleri takip etmektedir. Schwarzer testi sadece  $OR=9,01$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Harbord testi sadece  $OR \geq 4,02$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Thompson testinin yanlılığı tespit etme gücü  $OR=9,01$  ve örneklem hacmi küçük olduğu durum hariç her durumda %80'e ulaşmıştır. Genel olarak meta analizine alınan çalışma sayısı  $k=25$ 'ten  $k=50$ 'ye çıktığında sadece Egger testinin yanlılığı tespit etme gücü bakımından performansı değişmemiş, bu tez çalışmasında incelenen diğer testlerin ise yanlılığı tespit etme gücü bakımından performansları yükselmiştir.

Simülasyon sonucunda (Tablo 23–27) örneklem hacmi ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=100$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Begg, Egger ve Thompson testlerinin her koşulda tam olarak %80 güce ulaşmasıyla birlikte, en iyi performansı Begg, Egger ve Thompson testleri göstermiştir. Begg, Egger ve Thompson testlerini, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer testi  $OR \geq 4,02$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Harbord testi  $OR \geq 2,34$  ve küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Meta analizine alınan çalışma sayısı  $k=50$ 'den  $k=100$ 'e çıktığında bu tez çalışmasında incelenen Thompson, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Begg ve Egger testlerinin ise değişmemiştir.

Simülasyon sonucunda (Tablo 23–27) örneklem hacmi ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve meta analizine alınan çalışma sayısı  $k=150$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen Begg, Egger ve Thompson testlerinin her koşulda tam olarak %80 güce ulaşmasıyla birlikte, en iyi performansı Begg, Egger ve Thompson testleri göstermiştir. Begg, Egger ve Thompson testlerini, Harbord ve Schwarzer testleri takip etmektedir. OR değerinin artmasıyla ve küçük örneklem hacimlerinde Harbord testinin performansı yükselmiştir. Schwarzer testi küçük örneklem hacimlerinde %80 güce ulaşabilmiştir. Genel olarak meta analizine alınan çalışma sayısı  $k=100$ 'den  $k=150$ 'ye çıktığında bu tez çalışmasında incelenen bütün testlerin yanlılığı tespit etme gücü bakımından performansları değişmemiştir.

Simülasyon sonucunda (Tablo 23–27) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve örneklem hacmi  $n=8$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=1,50$  değerine kadar yükseliş eğiliminde iken bu değerden sonra düşmeye başlamıştır. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=2,34$  değerinden itibaren yükselmiştir. Egger testinin yanlılığı tespit etme gücü bakımından performansı ise  $OR=2,34$  değerine kadar yükselmiş, bu değerden sonra düşmüş ve daha sonra sabit kalmıştır.

Simülasyon sonucunda (Tablo 23–27) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve örneklem hacmi  $n=10$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Begg ve Egger testleri göstermiştir Begg ve Egger testlerini, Thompson, Harbord ve Schwarzer testleri takip etmektedir. Thompson testinin yanlılığı tespit etme gücü bakımından performansı  $OR= 2,34$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren düşmeye

başlamıştır. Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları  $OR=4,02$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren düşmeye başlamıştır. Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=4,02$  değerinden itibaren yükselmiştir. Genel olarak örnekleme alınan birim sayısının  $n=8$ 'den,  $n=10$ 'a çıkmasıyla Begg testinin yanlılığı tespit etme gücü bakımından performansı yükselirken, Harbord ve Schwarzer testlerinin düşmüş, Egger ve Thompson testlerinin ise değişmemiştir.

Simülasyon sonucunda (Tablo 23–27) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve örneklem hacmi  $n=12$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Schwarzer testinin yanlılığı tespit etme gücü bakımından performansı  $OR=4,02$  değerine kadar OR değerinin artmasıyla düşmüş,  $OR=4,02$  değerinden sonra yükselmiştir. Harbord testinin yanlılığı tespit etme gücü bakımından performansı ise  $OR=4,02$  değerinden sonra yükselmiştir. Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları ise  $OR=4,02$  değerine kadar yükseliş eğiliminde iken, bu değerden itibaren düşmeye başlamıştır. Genel olarak örnekleme alınan birim sayısının  $n=10$ 'dan,  $n=12$ 'ye çıkmasıyla Begg, Egger ve Harbord ve testlerinin testinin yanlılığı tespit etme gücü bakımından performansı düşmüş, Schwarzer testinin performansı değişmemiş, Thompson testinin performansı ise yükselmiştir.

Simülasyon sonucunda (Tablo 23–27) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve örneklem hacmi  $n=15$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg testinin yanlılığı tespit etme gücü bakımından performansı  $OR=1,50$  değerine kadar yükseliş eğiliminde iken bu değerden sonra düşme eğilimi göstermiştir. Thompson testinin yanlılığı tespit etme



gücü bakımından performansı  $OR=4,02$  değerine kadar, Egger testinin yanlılığı tespit etme gücü bakımından performansı ise  $OR=9,01$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren düşme eğilimi göstermiştir. Schwarzer testinin yanlılığı tespit etme gücü bakımından performansı  $OR$  değerinin artmasıyla düşmüştür.  $OR$  değerinin artmasıyla Harbord testinin yanlılığı tespit etme gücü bakımından performansı ise  $OR=4,02$  değerinden sonra yükselmiştir. Genel olarak örnekleme alınan birim sayısının  $n=12$ 'den,  $n=15$ 'e çıkmasıyla Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Schwarzer ve Harbord testlerinin performansları düşmüş, Thompson testinin performansı değişmemiştir.

Simülasyon sonucunda (Tablo 23–27) meta analizine alınan çalışma sayısı ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve örneklem hacmi  $n=20$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Begg testinin yanlılığı tespit etme gücü bakımından performansı  $OR=1,50$  değerine kadar yükseliş eğiliminde iken bu değerden sonra düşme eğilimi göstermiş ve daha sonra değişmemiştir. Thompson testinin yanlılığı tespit etme gücü bakımından performansı  $OR=9,01$  değerine kadar yükseliş eğiliminde iken bu değerden itibaren performansı düşmeye başlamıştır. Egger testinin yanlılığı tespit etme gücü bakımından performansı  $OR=4,02$  değerine kadar yükselme eğiliminde iken bu değerden sonra performansı düşmeye başlamıştır. Harbord testinin yanlılığı tespit etme gücü bakımından performansı ise  $OR=9,01$  değerinden itibaren yükselmiştir. Genel olarak örnekleme alınan birim sayısının  $n=15$ 'ten,  $n=20$ 'ye çıkmasıyla Begg ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselmiş, Schwarzer ve Harbord testlerinin performansları düşmüş, Egger testinin performansı değişmemiştir.

Simülasyon sonucunda (Tablo 23–27) meta analizine alınan çalışma sayısı ve  $OR$  değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve örneklem hacmi  $n=25$  olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam

olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson ve Harbord testi takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren test, hiçbir durumda %80 güce ulaşamayan Schwarzer testidir. OR değeri arttıkça Egger testinin yanlılığı tespit etme gücü bakımından performansı yükselme eğilimindedir. Harbord testinin yanlılığı tespit etme gücü bakımından performansı ise OR=9,01 değerinden itibaren yükselmiştir. Begg testinin performansı OR=1,50 değerine kadar yükseliş göstermiş, bu değerden sonra OR=9,01 değerine kadar düşme eğilimi göstermiş ve OR=9,01 değerinde tekrar yükselmiştir. Thompson testinin performansı OR=9,01 değerine kadar yükselme eğilimi göstermiş ve OR=9,01 değerinde tekrar düşmüştür. Genel olarak örnekleme alınan birim sayısının n=20'den, n=25'e çıkmasıyla Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Harbord ve Thompson testlerinin performansı düşmüş, Schwarzer testinin performansı ise değişmemiştir.

Simülasyon sonucunda (Tablo 23–27) meta analizine alınan çalışma sayısı ve OR değerleri ayrımı yapılmaksızın meta analizinde %60 yanlılık oluşturulduğu ve örneklem hacmi n=50 olduğu durumda genel olarak testler güçlerine göre değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Begg ve Egger testleri göstermiştir. Begg ve Egger testlerini, Thompson testi takip etmektedir. Yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testler, hiçbir durumda %80 güce ulaşamayan Harbord ve Schwarzer testleridir. Begg testinin performansı OR=1,50 değerine kadar yükseliş göstermiş, bu değerden sonra OR=9,01 değerine kadar düşme eğilimi göstermiş ve OR=9,01 değerinde tekrar yükselmiştir. OR değeri arttıkça Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğilimindedir. Egger testi OR=9,01 olduğu her durumda %80 güce ulaşmıştır. Genel olarak örnekleme alınan birim sayısının n=25'ten, n=50'ye çıkmasıyla Egger ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansları düşmüş, Thompson testinin performansı yükselmiş, Begg ve Schwarzer testlerinin performansı ise değişmemiştir.

Meta analizinde %60 yanlılık oluşturulduğu durum dikkate alınarak yapılan simülasyon senaryoları (Tablo 23–27) incelendiğinde örneklem hacmi, meta analizine

alınan çalışma sayıları ve OR değeri ayrımı yapılmaksızın genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. OR değerinin artmasıyla Begg ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları OR=2,34 değerine kadar yükselmiş, bu değerden itibaren düşmüştür. OR değerinin artmasıyla Egger testinin yanlılığı tespit etme gücü bakımından performansı OR=4,02 değerine kadar yükselmiş, bu OR değerinden itibaren düşmeye başlamıştır. Harbord testinin yanlılığı tespit etme gücü bakımından performansı, OR değerinin artmasıyla OR=2,34 değerinden itibaren, örneklem hacminin azalması ve meta analizine alınan çalışma sayılarının artması ile yükselmiştir. Schwarzer testinin yanlılığı tespit etme gücü bakımından performansı OR değerinin artması ile önce düşmüş, özellikle OR=4,02 değerinden itibaren, örneklem hacminin azalması ve meta analizine alınan çalışma sayılarının artması ile yükselmiştir. Bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları meta analizine alınan çalışma sayıları arttıkça yükselmiştir. Meta analizine alınan çalışma sayısı az iken bu tez çalışmasında kullanılan testlerin hiçbiri yanlılığı tespit etme gücü bakımından iyi performans göstermezken, meta analizine alınan çalışma sayısı fazla ise Begg, Egger ve Thompson testleri çok iyi performans göstermiştir. Meta analizine alınan çalışma sayısı  $k=20,25,50$  olduğu durumlarda da Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları iyidir. Örneklem hacmi arttıkça Begg, Egger ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansları yükselme eğiliminde iken, Schwarzer ve Harbord testlerinin performansları düşme eğilimindedir.

Simülasyon sonucunda (Tablo 8, 13, 18, 23) örneklem hacmi, meta analizine alınan çalışma sayısı, yanlılık derecesi ayrımı yapılmaksızın ve OR=1,00 olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi, daha sonra Thompson ve Schwarzer

testleri takip etmektedir. Harbord testi %80 güce ulaşamamıştır. Yanlılık derecesinin artmasıyla Begg, Egger, Schwarzer ve Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğilimindedir.

Simülasyon sonucunda (Tablo 9, 14, 19, 24) örneklem hacmi, meta analizine alınan çalışma sayısı, yanlılık derecesi ayırımı yapılmaksızın ve  $OR=1,50$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Begg testi göstermiştir. Begg testini, yakın bir performansla Egger testi, daha sonra Thompson testi takip etmektedir. Harbord ve Schwarzer testleri %80 güce ulaşamamıştır. Yanlılık derecesinin artmasıyla Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğiliminde, Thompson testinin ise %20 yanlılık oluşturulduğu duruma kadar yükselmiş sonrasında düşmüştür.

Simülasyon sonucunda (Tablo 10, 15, 20, 25) örneklem hacmi, meta analizine alınan çalışma sayısı, yanlılık derecesi ayırımı yapılmaksızın ve  $OR=2,34$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi, daha sonra Thompson testi ve Harbord testi takip etmektedir. Schwarzer testi %80 güce ulaşamamıştır. Yanlılık derecesinin artmasıyla Begg, Egger ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Thompson testinin ise düşmüştür.

Simülasyon sonucunda (Tablo 11, 16, 21, 26) örneklem hacmi, meta analizine alınan çalışma sayısı, yanlılık derecesi ayırımı yapılmaksızın ve  $OR=4,02$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi, daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Thompson testinin ise düşmüştür. Yanlılık derecesinin artmasıyla Begg ve Egger

testlerinin yanlılığı tespit etme gücü bakımından performansı ise %20 yanlılık oluşturulduğu duruma kadar yükselmiş, bu değerden sonra düşme eğilimine girmiştir.

Simülasyon sonucunda (Tablo 12, 17, 22, 27) örneklem hacmi, meta analizine alınan çalışma sayısı, yanlılık derecesi ayrımı yapılmaksızın ve  $OR=9,01$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi, daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Harbord testinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Egger ve Thompson testlerinin ise düşmüştür. Yanlılık derecesinin artmasıyla Begg testinin yanlılığı tespit etme gücü bakımından performansı ise %20 yanlılık oluşturulduğu duruma kadar yükselme, bu değerden sonra düşme eğilimindedir. Schwarzer testinin yanlılığı tespit etme gücü bakımından performansı ise yanlılık derecesinin artmasıyla önce artmış, sonra azalmış sonra yeniden artmıştır.

Simülasyon sonucunda (Tablo 8-27) örneklem hacmi, OR değerleri, yanlılık derecesi ayrımı yapılmaksızın ve meta analizine alınan çalışma sayısı  $k=8$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, Egger testi çok iyi performans göstermemekle birlikte en iyi performansı göstermiş, Egger testi hariç diğer testler %80 güce ulaşamamıştır. Egger testinin yanlılığı tespit etme gücü bakımından performansı %60 yanlılık oluşturulduğu durumda düşmüştür.

Simülasyon sonucunda (Tablo 8-27) örneklem hacmi, OR değerleri, yanlılık derecesi ayrımı yapılmaksızın ve meta analizine alınan çalışma sayısı  $k=10$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi takip etmektedir. Schwarzer, Harbord ve Thompson testleri %80 güce ulaşamayarak yanlılığı tespit etme gücü yönünden en kötü performansı gösteren testlerdir. Yanlılık derecesinin artmasıyla Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiştir.

Simülasyon sonucunda (Tablo 8-27) örneklem hacmi, OR değerleri, yanlışlık derecesi ayrımı yapılmaksızın ve meta analizine alınan çalışma sayısı  $k=15$  olduğu durumda genel olarak testlerin yanlışlığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi, daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlışlık derecesinin artmasıyla Begg, Egger, Harbord ve Schwarzer testlerinin yanlışlığı tespit etme gücü bakımından performansı yükselme eğilimindedir. Thompson testi ise sadece %40 yanlışlık oluşturulduğu durumda çok da iyi olmayan bir performans göstermiştir.

Simülasyon sonucunda (Tablo 8-27) örneklem hacmi, OR değerleri, yanlışlık derecesi ayrımı yapılmaksızın ve meta analizine alınan çalışma sayısı  $k=20$  olduğu durumda genel olarak testlerin yanlışlığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi, daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlışlık derecesinin artmasıyla Begg, Egger, Harbord ve Schwarzer testlerinin yanlışlığı tespit etme gücü bakımından performansı yükselme eğiliminde iken, Thompson testinin ise düşme eğilimindedir.

Simülasyon sonucunda (Tablo 8-27) örneklem hacmi, OR değerleri, yanlışlık derecesi ayrımı yapılmaksızın ve meta analizine alınan çalışma sayısı  $k=25$  olduğu durumda genel olarak testlerin yanlışlığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi, daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlışlık derecesinin artmasıyla Begg, Egger, Harbord ve Schwarzer testlerinin yanlışlığı tespit etme gücü bakımından performansı yükselme eğiliminde iken, Thompson testinin ise düşme eğilimindedir. Egger testi %40 ve %60 yanlışlık oluşturulduğu her durumda %80 güce ulaşmıştır.

Simülasyon sonucunda (Tablo 8-27) örneklem hacmi, OR değerleri, yanlışlık derecesi ayrımı yapılmaksızın ve meta analizine alınan çalışma sayısı  $k=50$  olduğu

durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Begg ve Egger testleri göstermiştir. Egger testini, Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg, Egger ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğiliminde iken, Schwarzer testinin değişmemiş, Thompson testinin ise yanlılık oluşturma derecesi %20 olduğu duruma kadar yükselme, bu değerden sonra düşme eğilimindedir. Begg ve Egger testleri %20, %40 ve %60 yanlılık oluşturulduğu her durumda %80 güce ulaşmıştır.

Simülasyon sonucunda (Tablo 8-27) örneklem hacmi, OR değerleri, yanlılık derecesi ayrımı yapılmaksızın ve meta analizine alınan çalışma sayısı  $k=100$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, en iyi performansı Egger ve Thompson testleri göstermiştir. Egger ve Thompson testlerini, yakın bir performansla Begg testi, daha sonra Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğiliminde iken, Egger ve Thompson testlerinin değişmemiştir. Egger ve Thompson testleri her yanlılık derecesinde %80 güce ulaşmış, Begg testi ise %20, %40 ve %60 yanlılık oluşturulduğu her durumda %80 güce ulaşmıştır.

Simülasyon sonucunda (Tablo 8-27) örneklem hacmi, OR değerleri, yanlılık derecesi ayrımı yapılmaksızın ve meta analizine alınan çalışma sayısı  $k=150$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, en iyi performansı Begg, Egger ve Thompson testleri göstermiştir. Begg, Egger ve Thompson testlerini Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğiliminde iken, Begg, Egger ve Thompson testlerinin değişmemiştir. Begg, Egger ve Thompson testleri her yanlılık derecesinde %80 güce ulaşmıştır.

Simülasyon sonucunda (Tablo 8-27) meta analizine alınan çalışma sayısı ve OR değerleri yanlılık derecesi ayrımı yapılmaksızın ve örneklem hacmi  $n=8$  olduğu

durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg, Egger, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğiliminde iken, Thompson testinin ise önce yükselmiş, sonra düşmüştür.

Simülasyon sonucunda (Tablo 8-27) meta analizine alınan çalışma sayısı ve OR değerleri yanlılık derecesi ayrımı yapılmaksızın ve örneklem hacmi  $n=10$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg, Egger, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğiliminde iken, Thompson testinin ise %20 yanlılık oluşturulduğu duruma kadar yükselmiş, sonra düşmüştür.

Simülasyon sonucunda (Tablo 8-27) meta analizine alınan çalışma sayısı ve OR değerleri yanlılık derecesi ayrımı yapılmaksızın ve örneklem hacmi  $n=12$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson testi, Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg, Egger, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiş, Thompson testinin ise düşmüştür.

Simülasyon sonucunda (Tablo 8-27) meta analizine alınan çalışma sayısı ve OR değerleri yanlılık derecesi ayrımı yapılmaksızın ve örneklem hacmi  $n=15$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir.



Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg, Egger, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğilimindedir, Thompson testinin ise düşmüştür.

Simülasyon sonucunda (Tablo 8-27) meta analizine alınan çalışma sayısı ve OR değerleri yanlılık derecesi ayırımı yapılmaksızın ve örneklem hacmi  $n=20$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg, Egger, Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğilimindedir, Thompson testinin ise düşmüştür.

Simülasyon sonucunda (Tablo 8-27) meta analizine alınan çalışma sayısı ve OR değerleri yanlılık derecesi ayırımı yapılmaksızın ve örneklem hacmi  $n=25$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve %80 güce ulaşamayan Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg, Egger ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı yükselme eğiliminde iken, Schwarzer testinin değişmemiş, Thompson testinin ise %20 yanlılık oluşturulduğu duruma kadar yükselmiş, sonra düşmüştür.

Simülasyon sonucunda (Tablo 8-27) meta analizine alınan çalışma sayısı ve OR değerleri yanlılık derecesi ayırımı yapılmaksızın ve örneklem hacmi  $n=50$  olduğu durumda genel olarak testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bu tez çalışmasında incelenen bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson testi ve %80 güce ulaşamayan Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından

performansı yükselme eğiliminde iken, Schwarzer ve Harbord testlerinin değişmemiş, Thompson testinin ise düşmüştür.

Yapılan simülasyon senaryoları (Tablo 8-27) incelendiğinde yanlılık derecesi, örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayrımı yapılmaksızın genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, bütün testlerin her koşulda tam olarak %80 güce ulaşamamasıyla birlikte, en iyi performansı Egger testi göstermiştir. Egger testini, yakın bir performansla Begg testi ve daha sonra Thompson, Harbord ve Schwarzer testleri takip etmektedir. Yanlılık derecesinin artmasıyla Begg, Egger ve Harbord testlerinin yanlılığı tespit etme gücü bakımından performansı yükselmiştir. Thompson testinin yanlılığı tespit etme gücü bakımından performansı yanlılık derecesi arttıkça %40 yanlılık oluşturulduğu duruma kadar yükselmiş, bu yanlılık derecesinden itibaren düşmeye başlamıştır. Genel olarak meta analizine alınan çalışma sayısının artışı ile bu tez çalışmasında incelenen bütün testlerin yanlılığı tespit etme gücü bakımından performansları yükseliş eğilimindedir. Genel olarak örneklem hacminin artması ile Harbord ve Schwarzer testlerinin yanlılığı tespit etme gücü bakımından performansları düşmüş, Thompson testinin performansı yükselmiştir. Begg ve Egger testlerinin yanlılığı tespit etme gücü bakımından performansları  $n=25$  değerine kadar genel olarak yükselme eğiliminde iken, bu değerden sonra düşmeye başlamıştır. Genel olarak testlerin OR değerinin artışı ile yanlılığı tespit etme gücü bakımından performansları incelendiğinde Harbord testinin yükselmiş, Egger ve Thompson testlerinin  $OR=2,34$  değerine kadar, Begg testinin  $OR=1,50$  değerine kadar yükselmiş daha sonra düşmüştür. OR değerinin artışı ile Schwarzer testinin yanlılığı tespit etme gücü bakımından performansı  $OR=1,50$  değerine kadar düşmüş, daha sonra OR değerinin artışı ile tekrar yükselmiştir.

## 5. TARTIŞMA ve SONUÇ

Meta analizinin en önemli adımlarından bir tanesi, meta analizine dahil edilen çalışmaların ilgili konu ile alakalı bütün çalışmaların bulunduğu anakütleyi temsil edip etmediğini, dolayısıyla meta analizi sonucu elde edilen bilgilerin yanlı olup olmadığını belirlemektir. Meta analizinde yanlılığın çalışma sonuçlarını etkilemeyecek düzeyde olduğu ortaya konulmalıdır. Meta analizinde yanlılık belirlemede öznel bir değerlendirme sunan funnel grafiklerinin objektif olarak değerlendirilmesini, daha temkinli yorumlanmasının kolaylaştırılmasını sağlayan istatistiksel testler, meta analizi sonuçlarının güvenilirliğini değerlendirmemize imkân sunar. Yanlılığı tespit etmeye yönelik istatistiksel testlerin performansının incelenmesi, ilgili istatistiksel testlerin ne zaman ve hangi koşullarda kullanılacağı konusunda bilgi sağlar. Funnel grafiğine dayalı testler, basitlikleri ve kullanıcı dostu yazılımların kullanılabilirliği nedeniyle yaygın olarak kullanılmaktadır. Fakat bu testler, bir meta analizinin sonuçlarının yorumlanmasında dikkat edilmesi gereken bir sorun konusunda sadece uyarıda bulunur, soruna bir çözüm getirmezler. Funnel grafiğine dayalı testlerin kullanımının artması, yanlılığın tespitinin raporlanması, daha doğru sonuçların çıkarılmasını sağlayacak, meta analizi metodolojisine de katkıda bulunacaktır.

Bu tez çalışmasında, ikili değer alan verilerin meta analizinde funnel grafiği asimetri değerlendirmesinde yanlılık belirlemek için kullanılan literatürde yer alan Begg, Egger, Thompson, Schwarzer ve Harbord testlerinin Tip-I hata oranları ve güçleri yönünden performanslarının farklı çalışma sayıları, farklı örneklem hacimleri, farklı yanlılık dereceleri ve farklı hastalık-sonuç oranları durumlarında nasıl etkilendiğinin incelenmesi amaçlanmıştır.

Meta analizinde yanlılığın olmadığı durum dikkate alınarak yapılan simülasyon senaryoları incelendiğinde örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayırımı yapılmaksızın genel olarak Tip-I hata oranları nominal

değere göre değerlendirildiğinde en iyi performansı Schwarzer testi göstermiştir. Schwarzer testinin Tip-I hatayı korumaya yönelik performansı genel olarak meta analizine alınan çalışma sayılarının artması ile düşmüş, örnekleme alınan birim sayısının artması ile yükselmiştir. OR değerinin artması, örneklem hacminin azalması ve meta analizine alınan çalışma sayılarının artması ile Schwarzer testinin Tip-I hata oranlarının ve dolayısıyla nominal değerden sapmanın oldukça arttığı görülmektedir. Simülasyon senaryoları incelendiğinde yanlılık derecesi, örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayırımı yapılmaksızın genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde en kötü performansı Schwarzer testi göstermiştir. Schwarzer testinin yanlılığı tespit etme gücü bakımından performansı genel olarak meta analizine alınan çalışma sayısının artışı ile yükselmiş, örneklem hacminin artması ile düşmüştür. OR değerinin artışı ile Schwarzer testinin yanlılığı tespit etme gücü bakımından performansının yükseldiği söylenebilir. Jin ve arkadaşları (2014) ve Schwarzer ve arkadaşları (2007) çalışmamıza benzer şekilde Schwarzer testinin uygun Tip-I hata oranlarına sahip olduğunu, ancak testin gücünün düşük ve yanlılığı en az tespit eden test olduğunu göstermiştir. Ayrıca Jin ve arkadaşları da (2014) bizim çalışmamızda olduğu gibi meta analizine alınan çalışma sayısının artması ile yöntemin gücünün arttığını bulmuştur. Bu tez çalışmasında ele alınan simülasyon senaryolarında meta analizinde %60 yanlılık oluşturulduğu durumda OR değeri 1'e yakın ve  $k \geq 100$  ise Schwarzer testinin gücü %50'nin üzerine çıkmakta, aynı zamanda OR değeri 1'e yakın ve  $k \geq 100$  ise uygun Tip-I hata oranları verdiği için, Schwarzer testi bu koşullar için önerilebilir.

Meta analizinde yanlılığın olmadığı durum dikkate alınarak yapılan simülasyon senaryoları incelendiğinde örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayırımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, Schwarzer testinden sonra en iyi performansı gösteren test Harbord testidir. Harbord testinin Tip-I hatayı korumaya yönelik performansı genel olarak meta analizine alınan çalışma sayılarının artması ile düşmüş, örnekleme alınan birim sayısının artması ile yükselmiştir. OR değerinin artması, örneklem hacminin azalması ve meta analizine alınan çalışma sayılarının artması ile Harbord testinin Tip-I hata oranlarının ve dolayısıyla nominal değerden sapmanın

oldukça arttığı görülmektedir. Harbord ve arkadaşları (2006)'nın ve Jin ve arkadaşlarının (2014) yapmış oldukları çalışmalarda, bizim çalışmamıza benzer şekilde, çalışmalar arası heterojenlik çok az olduğunda veya hiç olmadığına, Harbord testinin, nominal düzeye yakın Tip-I hata oranlarına sahip olduğu bulunmuştur. Simülasyon senaryoları incelendiğinde yanlılık derecesi, örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayırımı yapılmaksızın genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde Harbord testinin performansı çok iyi olmamakla birlikte, Schwarzer testinden daha iyidir. Harbord testinin yanlılığı tespit etme gücü bakımından performansı genel olarak meta analizine alınan çalışma sayısının, yanlılık derecesinin, OR değerinin artışı ile yükselmiş, örneklem hacminin artışı ile düşmüştür. Jin ve arkadaşları da (2014) bizim çalışmamızda olduğu gibi meta analizine alınan çalışma sayısının artması ile yöntemin gücünün arttığını bulmuştur. Sterne ve arkadaşları (2008) ikili sonuçlar için heterojenliğin olmadığı durumda Harbord testinin kullanılabileceğini belirtmiştir.

Meta analizinde yanlılığın olmadığı durum dikkate alınarak yapılan simülasyon senaryoları incelendiğinde örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayırımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, Thompson testi sadece  $k=8$  gibi meta analizine alınan çalışma sayısı az olduğu birkaç durumda Tip-I hatayı koruma yönünde performans göstermiştir. Thompson testinin Tip-I hata oranları genel olarak meta analizine alınan çalışma sayısı ve örnekleme alınan birim sayısının artması ile yükselmiştir. Thompson testi her ne kadar iyi performans göstermese de meta analizine alınan çalışma sayısı ve örneklem hacmi az olduğunda Tip-I hata oranları, nominal değere daha yakındır. OR değerinin artmasıyla Thompson testinde Tip-I hata oranlarının ve dolayısıyla nominal değerden sapmanın önce arttığı, sonra azaldığı saptanmıştır. Simülasyon senaryoları incelendiğinde yanlılık derecesi, örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayırımı yapılmaksızın genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, Thompson testinin performansı, Harbord ve dolayısı ile Schwarzer testinin performansından daha iyidir. Thompson testinin yanlılığı tespit etme gücü bakımından performansı yanlılık derecesi arttıkça %40

yanlılık oluşturulduğu duruma kadar yükselmiş, bu yanlılık derecesinden itibaren düşmeye başlamıştır. Thompson testinin yanlılığı tespit etme gücü bakımından performansı meta analizine alınan çalışma sayısının ve örneklem hacminin artışı ile yükselmiştir. OR değerinin artışı ile Thompson testlerinin yanlılığı tespit etme gücü bakımından performansı belirli bir değere kadar yükselme eğiliminde iken sonra düşmüştür.

Meta analizinde yanlılığın olmadığı durum dikkate alınarak yapılan simülasyon senaryoları incelendiğinde örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayrımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, Egger testi Tip-I hatayı nominal düzeyde koruma yönünde başarısız olmuş ve oldukça sapmalı sonuçlar vermiştir. Egger testinin genel olarak meta analizine alınan çalışma sayılarının ve örnekleme alınan birim sayısının artması ile Tip-I hata oranları artmıştır. Egger testi her ne kadar iyi performans göstermese de meta analizine alınan çalışma sayısı ve örneklem hacmi az olduğunda Tip-I hata oranları, nominal değere daha yakındır. OR değerinin ve meta analizine alınan çalışma sayısının artması ile Egger testinde Tip-I hata oranlarının ve dolayısıyla nominal değerden sapmanın arttığı saptanmıştır. Schwarzer ve arkadaşları (2007)'nin çalışma sonuçlarına benzer şekilde biz de etki büyüklüğünün artmasıyla Tip-I hata oranlarının şiştiğini görmekteyiz. Ayrıca Sterne ve arkadaşlarının (2000) çalışmasında da büyük etki büyüklükleri, meta analizindeki tüm çalışmaların benzer örneklem hacimlerine sahip olması durumlarında anormal yükseklikte, nominal değeri büyük ölçüde aşan Tip-I hata oranları ortaya çıktığı ifade edilmiştir. Simülasyon senaryoları incelendiğinde yanlılık derecesi, örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayrımı yapılmaksızın genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, en iyi performansı Egger testi göstermiştir. Yanlılık derecesinin, meta analizine alınan çalışma sayısının artmasıyla Egger testinin yanlılığı tespit etme gücü bakımından performansı yükselmiştir. Örneklem hacminin artması ile Egger testinin yanlılığı tespit etme gücü bakımından performansı n=25 değerine kadar yükselme eğiliminde iken, bu değerden sonra düşmeye başlamıştır. OR değerinin artışı ile Egger testinin yanlılığı tespit etme gücü bakımından performansı belirli bir OR değerine kadar yükselme eğiliminde iken sonra düşmüştür. Peters ve arkadaşları (2006), Jin ve arkadaşları

(2014), Schwarzer ve arkadaşları (2007), Macaskill ve arkadaşları (2001) Egger testinin, yüksek güce sahip olduğunu, çalışma sayısının artması ile yöntemin gücünün arttığını, ancak uygun olmayan Tip-I hata oranına sahip olduğunu, etki büyüklüğü ve çalışma sayısı arttıkça Tip-I hata oranının nominal değeri aştığını bulmuştur. Kale ve Nirpharake (2017), Sterne ve arkadaşları (2000) çalışmamıza benzer olarak yanlılık derecesi, meta analizine alınan çalışma sayısı ve OR değeri arttıkça gücün arttığını, 10 veya daha az çalışma sayısına dayanan veya ciddi yanlılığın olmadığı meta analizleri için gücün daha düşük olduğunu bulmuştur.

Meta analizinde yanlılığın olmadığı durum dikkate alınarak yapılan simülasyon senaryoları incelendiğinde örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayrımı yapılmaksızın genel olarak Tip-I hata oranları nominal değere göre değerlendirildiğinde, Begg testi Tip-I hatayı nominal düzeyde koruma yönünde başarısız olmuş ve oldukça sapmalı sonuçlar vermiştir. Begg testinin genel olarak meta analizine alınan çalışma sayısının ve örnekleme alınan birim sayısının artması ile Tip-I hata oranları artmıştır. Begg testi her ne kadar iyi performans göstermese de meta analizine alınan çalışma sayısı ve örneklem hacmi az olduğunda Tip-I hata oranları, nominal değere daha yakındır. OR değerinin ve meta analizine alınan çalışma sayısının artması ile Begg testinde Tip-I hata oranlarının ve dolayısıyla nominal değerden sapmanın oldukça arttığı saptanmıştır. Schwarzer ve arkadaşları (2007) çalışmamıza benzer olarak etki büyüklüğünün artmasıyla Tip-I hata oranlarının şiştiğini belirtmiştir. Ayrıca Sterne ve arkadaşlarının (2000) çalışmasında da büyük etki büyüklükleri, meta analizindeki tüm çalışmaların benzer örneklem hacimlerine sahip olması durumlarında anormal yükseklikte, nominal değeri büyük ölçüde aşan Tip-I hata oranları ortaya çıktığı ifade edilmiştir. Yanlılık derecesi, örneklem hacmi, meta analizine alınan çalışma sayıları ve OR değeri ayrımı yapılmaksızın oluşturulan simülasyon senaryoları incelenerek, genel olarak bu tez çalışmasında kullanılan testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde Begg testi, Egger testine yakın ve iyi bir performans göstermiştir. Yanlılık derecesinin, meta analizine alınan çalışma sayısının artmasıyla Begg testinin yanlılığı tespit etme gücü bakımından performansı yükselmiştir. Genel olarak örneklem hacminin artması ile Begg testinin yanlılığı tespit etme gücü bakımından performansı  $n=25$  değerine kadar yükselme eğiliminde iken, bu değerden sonra düşmeye başlamıştır. OR

değerinin artışı ile Begg testinin yanlılığı tespit etme gücü bakımından performansı belirli bir değere kadar yükselmiş sonra düşmüştür. Bizim çalışmamızda olduğu gibi Kale ve Nirpharake (2017) ve Jin ve arkadaşlarının (2014) çalışmasında da meta analizine alınan çalışma sayısının artması ile yöntemin gücünün arttığı bulunmuştur. Sterne ve arkadaşları (2000) çalışmamıza benzer olarak yanlılık, meta analizine alınan çalışma sayısı ve OR değeri arttıkça gücün arttığını, 10 veya daha az çalışma sayısına dayanan veya ciddi yanlılığın olmadığı meta analizleri için gücün daha düşük olduğunu bulmuştur. Begg ve Mazumdar (1994), yanlılığın güçlü olduğu durumu ele almadıklarını, 75 çalışma içeren büyük meta analizleri için testin güçlü, 25 çalışma içeren meta analizler için gücün ortaya düzeyde olduğunu, bu nedenle test istatistiksel olarak anlamlı değilse yanlılığın göz ardı edilmemesi gerektiğini belirtmiştir. Simülasyon senaryolarımız neticesinde meta analizinde %60 yanlılık oluşturulduğu durumda OR değeri 1'e yakın ve  $k \geq 8$  ise Begg testinin, gücü %60'ların üzerine çıkmakta, aynı zamanda Begg testinin Tip-I hata oranları Peterson kriterlerine uymasa da nominal değere yaklaşmakta olduğundan OR değeri 1'e yakın ve  $k \geq 8$  ise, Begg testi bu koşullar için önerilebilir.

Bizim çalışmamızda olduğu gibi Jin ve arkadaşlarının (2014), Macaskill ve arkadaşlarının (2001), Sterne ve arkadaşlarının (2000) ve Begg ve Mazumdar'ın (1994) çalışmalarında meta analizine alınan çalışma sayısının artması ile yöntemlerin gücünün arttığı, meta analizine 10 veya daha az çalışma dahil edildiğinde, çalışmalarına dahil ettikleri tüm testlerinin gücünün düşük olduğu bulunmuştur. Ioannidis ve Trikalinos (2007), Sterne ve arkadaşları (2008) funnel grafiği asimetrisi için istatistiksel testlerin meta analizine alınan 10 veya daha fazla çalışma söz konusu olduğunda uygun olabileceğini önermektedir. Ayrıca, Schwarzer ve arkadaşları (2007) etki büyüklüğü ve bunun standart hatası arasındaki ilişkinin, Egger ve Begg testinde OR değerinin artması ile Tip-I hata oranlarının şişmesinin nedeni olarak açıklamaktadır. Yani yanlılığı belirlemek için kullanılan bu iki yöntemin kendisi yanlıdır ve yanlılığı olduğundan fazla tahmin eder denilebilir.

Funnel grafiği asimetrisini değerlendirmek için, Tip-I hata oranları ve gücü yüksek olan ve Tip-I hata oranları ve gücü düşük olan iki test birbirinin tamamlayıcısı olarak kullanılarak, iki testin birbiri ile uyumlu sonuçlar verdiği durumlarda yanlılık olduğu ya da olmadığı yönünde karar verilebilir. Yanlılık olduğu yönünde çıkan



sonuçlarda ya da iki test sonucu arasında uyumsuzluk olduğunda duyarlılık analizleri, meta analizine dahil edilmeyen çalışmaların kontrolü gibi seçenekler değerlendirilebilir.

Yapılan simülasyon çalışmaları sonucunda bu tez çalışmasında ele alınan testlerin hiçbiri senaryoların tamamında, belirlenen Tip-I hata olasılık değerini koruyamamasıyla birlikte güçlü de değildir. Ele alınan koşullarda, testler arasında her tür veri için kullanılabilir yanlılığı saptamak için en iyi test yoktur. Schwarzer ve Harbord testleri funnel grafiği asimetrisini tanımlamak için düşük istatistiksel güce, Begg, Egger ve Thompson testleri ise funnel grafiği asimetrisi olmadığında şişirilmiş Tip-I hata oranlarına sahiptir. Testlerin performansları meta analizine dahil edilen çalışmaların sayısı, yanlılık seviyeleri, hastalık-sonuç oranları ve örneklem hacmi ile değiştiğinden yanlılığı saptamaya yönelik yöntemler seçilirken bu değişkenlerin dikkate alınması gerekmektedir. Bu tez çalışmasında ele alınan simülasyon senaryoları sonucunda Tip-I hata olasılık değerini koruma yönünde en iyi performansı meta analizine alınan çalışma sayısı, örneklem hacmi az ise Schwarzer, fazla ise Harbord testlerinin gösterdiği söylenebilir. Simülasyon senaryoları sonucunda testlerin yanlılığı tespit etme gücü bakımından performansları değerlendirildiğinde, en iyi performansı Egger testi göstermiştir. Bu tez çalışmasında ele alınan testler değerlendirilirken, funnel grafiği asimetrisinin kanıtı sağlanmasa bile yanlılık göz ardı edilmemelidir.

## 6. KAYNAKLAR

- Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). "How to perform a meta-analysis with R: a practical tutorial." *Evidence-Based Mental Health*, 153–160. doi: 10.1136/ebmental-2019-300117.
- Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50(4), 1088-1101. Erişim adresi: <https://www.jstor.org/stable/2533446>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta Analysis*. United Kingdom, John Wiley & Sons, Ltd.
- Debray, T. P. A., Moons, K. G. M., & Riley, R. D. (2017). Detecting small-study effects and funnel plot asymmetry in meta-analysis of survival data: A comparison of new and existing tests. *WileyResSynMeth*, 9, 41–50. doi: 10.1002/jrsm.1266
- Deeks, J. J. (2002). Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine*, 21, 1575-1600. doi: 10.1002/sim.1188
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7, 177–188. Doi: 10.1016/0197-2456(86)90046-2
- Duval, S., & Tweedie, R. (2000). Trim and Fill: A simple funnel plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56, 455-463. doi: 10.1111/j.0006-341x.2000.00455.x
- Duval, S. J. (2005). The trim and fill method. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.). *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 127–144). Chichester, England: Wiley.
- Easterbrook, P. J., Berlin, J. A., Gopalan, R., & Matthews, D.R. (1991). Publication bias in clinical research. *Lancet*. 337, 867–72. doi: 10.1016/0140-6736(91)90201-y
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629–634. doi: 10.1136/bmj.315.7109.629
- Harbord, R. M., Egger, M., & Sterne, J. A. C. (2006). A modified test for smallstudy effects in meta-analyses of controlled trials with binary endpoints. *Statistics in Medicine*, 25, 3443–3457. doi: 10.1002/sim.2380
- Harbord, R. M., Harris, R. J., & Sterne, J. A. C. (2009). Updated tests for small-study effects in meta-analyses. *The Stata Journal*, 9(2), 197–210. doi: 10.1177/1536867X0900900202
- Hardy, R. J., & Thompson, S. G. (1998). Detecting and Describing Heterogeneity in Meta Analysis. *Statistics in Medicine*, 17(8), 841-856. doi: 10.1002/(sici)1097-0258(19980430)17:8<841::aid-sim781>3.0.co;2-d
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying Heterogeneity in A Meta Analysis. *Statistics in Medicine*, 21(11), 1539-1558. doi: 10.1002/sim.1186
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring Inconsistency in Meta Analysis. *British Medical Journal*, 327, 557-560. doi: 10.1136/bmj.327.7414.557

- Higgins, J. P. T., & Green, S. (2006). *Cochrane Handbook for Systematic Review of Interventions* 4.2.6. In: The Cochrane Library. John Wiley & Sons, Ltd.
- Higgins, J. P. T., & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of Interventions*. A John Wiley & Sons, Ltd., Publication.
- Hsiung, T. H., & Olejnik, S. (1996). Type I Error Rates and Statistical Power for the James Second-Order Test and the Univariate F Test in Two-Way Fixed-Effects ANOVA Models under Heteroscedasticity and/or Nonnormality. *The Journal of Experimental Education*, 65(1), 57-71. Erişim adresi: <https://www.jstor.org/stable/20152506>
- Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistics or  $I^2$  Index? *Psychological Methods*, 11(2), 193-206. doi: 10.1037/1082-989X.11.2.193
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage Publications, Inc.
- Ioannidis, J. P. A. (1998). Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA*, 279(4), 281-286. doi: 10.1001/jama.279.4.281
- Ioannidis, J. P. A., & Trikalinos, T. A. (2007). The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *CMAJ*, 176(8), 1091–1096. doi: 10.1503/cmaj.060410
- Jennions, M. D., & Moller, A. P. (2002). *Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution*. Proceedings of the Royal Society of London Series B: Biological Sciences, 269(1486),43–48. doi: 10.1098/rspb.2001.1832
- Jin, Z. C., Zhou, X. H., & He, J. (2014). Statistical methods for dealing with publication bias in meta-analysis. *Statist. Med.* 34(2), 343–360. doi: 10.1002/sim.6342
- Kale, J., & Nirpharake, A. (2017). Comparison of Different Methods of Detecting Publication Bias. *PhUSE*. Paper AS04. Erişim adresi: <https://www.lexjansen.com/phuse/2017/as/AS04.pdf>
- Macaskill, P., Walter, S.D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, 20(4),641–654. doi: 10.1002/sim.698
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22(4), 719-748. doi: 10.1093/jnci/22.4.719
- Mittlböck, M., & Heinzl, H. (2006). A simulation study comparing properties of heterogeneity measures in meta analyses. *Statistics in Medicine*, 25(24), 4321-4333. doi: 10.1002/sim.2692
- Normand, S. L. T. (1999). Tutorial in Biostatistics Meta-Analysis: Formulating, Evaluating, Combining and Reporting. *Statistics in Medicine*, 18(3), 321-359. doi: 10.1002/(SICI)1097-0258(19990215)18:3<321::AID-SIM28>3.0.CO;2-P
- Peters, J. I., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2006). Comparison of Two Methods to Detect Publication Bias in Meta-analysis. *JAMA*, 295(6), 676-680. doi: 10.1001/jama.295.6.676
- Peterson K. (2002). Six Modifications Of The Aligned Rank Transform Test For Interaction. *Journal of Modern Applied Statistical Methods*, 1(1), 100-109. Doi: 10.22237/jmasm/1020255240

- Petitti, D. B. (2000). *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis*. Oxford University Press.
- Petitti, D. B. (2001). Approaches to Heterogeneity in Meta-Analysis. *Statistics in Medicine*, 20(23), 3625-3633. doi: 10.1002/sim.1091
- Rothstein, H. R., Sutton, A.J., & Borenstein, M. (2005). *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. John Wiley & Sons, Ltd.
- Rücker, G., Schwarzer, G., Carpenter, J. R., & Schumacher, M. (2008). Undue Reliance on  $I^2$  in Assessing Heterogeneity May Mislead. *BMC Med Res Methodol*, 8, 79. doi: 10.1186/1471-2288-8-79
- Schwarzer, G., Antes, G., & Schumacher, M. (2007). A Test for Publication Bias in Meta-Analysis With Sparse Binary Data. *Statistics in Medicine*, 26, 721-733, doi: 10.1002/sim.2588.
- Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-Analysis with R*. Springer International Publishing, Switzerland.
- Schwarzer, G., Carpenter, J. R., & Rücker, G. (2022). *Statistical Methods for Sensitivity Analysis in Meta Analysis*. Erişim Adresi: <https://link.springer.com/book/10.1007/978-3-319-21416-0>
- Song, F., Parekh, S., Hooper, L., Loke, Y. K., Ryder, J., Sutton, A.J., ...Harvey, I. (2010). Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess*, 14(8), doi: 10.3310/hta14080.
- Spector, T. D., & Thompson, S. G. (1991). The potential and limitations of meta-analysis. *Journal of Epidemiology and Community Health*, 45(2), 89-92. doi: 10.1136/jech.45.2.89
- Stern, J. M., & Simes, R. J. (1997). Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ*, 315, 640-645. doi: 10.1136/bmj.315.7109.640
- Sterne, J. A. C., Gavaghan, D., & Egger, M. (2000). Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology*, 53(11), 1119-1129. doi: 10.1016/s0895-4356(00)00242-0
- Sterne, J. A. C., & Harbord, R. M. (2004). Funnel plots in meta-analysis. *The Stata Journal*, 4 (2), 127-141. Erişim adresi: <http://www.stata-journal.com/sjpdf.html?articlenum=st0061>
- Sterne, J., Egger, M., & Moher, D. (2008). *Addressing reporting biases. Cochrane handbook for systematic reviews of intervention. version 5.0.0*. The Cochrane Collaboration. Chapter 10.
- Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., ..., Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal*, 342, 1-8, doi: 10.1136/bmj.d4002.
- Stevenson, M., & Sargean, E. (2022). *Tools for the Analysis of Epidemiological Data*. Erişim adresi: <https://cran.r-project.org/web/packages/epiR/epiR.pdf>
- Sutton, A. J., Duval, S. J., Tweedie, R. L., Abrams, K. R., & Jones, D. R. (2000). Empirical assessment of effect of publication bias on meta-analyses. *BMJ*, 320(7249), 1574-1577. doi: 10.1136/bmj.320.7249.1574
- Süt, N., & Şenocak, M. (2007). Relatif risk ölçütünün odds oranı, atfedilen risk ve tedaviye gerekli sayı ölçütleriyle karşılaştırılması. *Trakya Üniversitesi Tıp Fakültesi Dergisi*, 24(3), 213-221.

- Takkouche, B., Cadarso-Suarez, C. C., & Spiegelman, D. (1999). Evaluation of Old and New Tests of Heterogeneity in Epidemiologic Meta-Analysis. *American Journal of Epidemiology*, 150(2), 206-214. doi: 10.1093/oxfordjournals.aje.a009981
- Tang, J. L., & Liu, J. L. Y. (2000). Misleading funnel plot for detection of bias in meta-analysis. *Journal of Clinical Epidemiology*, 53(5), 477-484. doi: 10.1016/s0895-4356(99)00204-8
- Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine*, 18(20), 2693-2708. doi: 10.1002/(sici)1097-0258(19991030)18:20<2693::aid-sim235>3.0.co;2-v
- Toews, I., Booth, A., Berg, R. C., Lewin, S., Glenton, C., Munthe-Kaas, H. M., ..., Meerpohl, J. J. (2017). Further exploration of dissemination bias in qualitative research required to facilitate assessment within qualitative evidence syntheses. *Journal of Clinical Epidemiology*. 88, 133-139. doi: 10.1016/j.jclinepi.2017.04.010
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1-48. Erişim adresi: <https://doi.org/10.18637/jss.v036.i03>.
- Wang, K. S., & Liu, X. (2016). Statistical methods in the meta-analysis of prevalence of human diseases. *Journal of Biostatistics and Epidemiology*. 2(1), 20-24. Erişim adresi: <https://jbe.tums.ac.ir/index.php/jbe/article/view/46>
- Whitehead, A. (2002). *Meta-Analysis of Controlled Clinical Trials*. Wiley Publication, England.
- Wickham, H., François, R., & Henry, L. (2022). A Grammar of Data Manipulation. Erişim adresi: <https://cran.r-project.org/web/packages/dplyr/index.html>
- Wickham, H., & Hester, J. (2022). Read Rectangular Text Data. Erişim adresi: <https://github.com/tidyverse/readr>.

## 7. SİMGELER ve KISALTMALAR

$\chi^2$ : Ki-kare

Ç.A.: Çalışmalar arası

Ç.İ.: Çalışma içi

F: Sabit etki modeli

FIV: Tahminin ters varyansı

FPV: Birleştirilmiş varyansın tersi

G.A.: Güven aralığı

H: Hasta grubu

k: Meta analizine alınan çalışma sayısı

K: Kontrol grubu

MH: Mantel Haenszel yöntemi

n: Örneklem hacmi

NNH: Yan etkiye gerekli sayı (Number Needed To Harm)

NNT: Tedaviye gerekli sayı (The Number Needed To Treat)

OR: Odds oranı (Odds Ratio)

$OR_{D\&L}$ : DerSimonian Laird yöntemi elde edilen birleştirilmiş odds oranı

$OR_{MH}$ : Mantel Haenszel yönteminden elde edilen birleştirilmiş odds oranı

$OR_{Peto}$ : Peto yönteminden elde edilen birleştirilmiş odds oranı

P: Hastalık-Sonuç oranları

$P_H$ : Hasta grubundaki durum riski olasılığı

$P_K$ : Kontrol grubundaki durum riski olasılığı

R: Rasgele etkiler modeli

RD: Risk farkı (Risk Difference)

$RD_{GV}$ : Genel varyansa dayalı yöntem ile elde edilen birleştirilmiş risk farkı

RR: Relatif risk (Relative Risk)

$RR_{GV}$ : Genel varyansa dayalı yöntem ile elde edilen birleştirilmiş relatif risk

SE: Standart hata (Standard Error)

## 8. TEŞEKKÜR

Doktora eğitimimin ve tez çalışmamın her aşamasında yoğun çalışma temposu içerisinde zaman ayırarak desteğini, önerilerini ve değerli bilgilerini esirgemeyen, maddi ve manevi her zaman yanımda olan değerli hocam ve tez danışmanım sayın Prof. Dr. İlker ERCAN'a güveni, sabrı ve emekleri için en içten saygı ve teşekkürlerimi sunarım. Tez izleme komitesinde yer alarak bu tez çalışmasının gerçekleşmesine değerli katkılarını sunan değerli hocalarım Prof. Dr. Berna YAZICI'ya ve Prof. Dr. Gökhan OCAKOĞLU'na, doktora eğitimim boyunca eğitimime katkıda bulunan anabilim dalımızdaki hocalarıma şükranlarımı sunarım. Ayrıca her koşulda bana destek olan eşime, anneme, eşimin anne ve babasına, canım kızıma, arkadaşlarıma ve yaşadığı süre boyunca bana her koşulda destek olan sevgili rahmetli babama teşekkürü bir borç bilirim.

## 9. ÖZGEÇMİŞ

.....yılında .....’de doğan Fisun KAŞKIR KESİN, 2005 yılında başladığı Eskişehir Osmangazi Üniversitesi İstatistik bölümünden 2009 yılında bölüm birincisi olarak mezun olmuştur. Yüksek lisans eğitimini 2009-2012 yılları arasında Eskişehir Osmangazi Üniversitesi İstatistik Anabilim Dalı İstatistik Bilgi Sistemleri Bilim Dalında tamamlamıştır. 2015 yılında Uludağ Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalında doktora eğitimine başlamıştır. 2014 yılından bu yana Düzce Üniversitesi, Sosyal Bilimler Meslek Yüksekokulunda Öğretim Görevlisi olarak görev yapmaktadır.