

T.C.
MERSİN ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK VE TIBBİ BİLİŞİM
ANABİLİM DALI

**META ANALİZİNDE HETEROJENLİĞİN
SAPTANMASINDA KULLANILAN YÖNTEMLERİN
SİMÜLASYON TEKNİĞİ İLE KARŞILAŞTIRILMASI**

Semra ERDOĞAN
DOKTORA TEZİ

DANIŞMAN
Doç. Dr. E. Arzu KANIK

Tez No:15

MERSİN- 2011

T.C.
MERSİN ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK VE TIBBİ BİLİŞİM
ANABİLİM DALI

META ANALİZİNDE HETEROJENLİĞİN SAPTANMASINDA KULLANILAN YÖNTEMLERİN SİMÜLASYON TEKNİĞİ İLE KARŞILAŞTIRILMASI

Semra ERDOĞAN
DOKTORA TEZİ

DANIŞMAN
Doç. Dr. E.Arzu KANIK

Bu tez, Mersin Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından BAP-SBE BTB (SE) 2009-9 DR nolu proje olarak desteklenmiştir.

Tez No:15

MERSİN- 2011

T.C.
MERSİN ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK VE TIBBİ BİLİŞİM
ANABİLİM DALI

META ANALİZİNDE HETEROJENLİĞİN SAPTANMASINDA KULLANILAN YÖNTEMLERİN SİMÜLASYON TEKNİĞİ İLE KARŞILAŞTIRILMASI

Semra ERDOĞAN
DOKTORA TEZİ

DANIŞMAN
Doç. Dr. E.Arzu KANIK

Bu tez, Mersin Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından BAP-SBE BTB (SE) 2009-9 DR nolu proje olarak desteklenmiştir.

Tez No:15

MERSİN- 2011

Mersin Üniversitesi Sağlık Bilimleri Enstitüsü

Biyoistatistik ve Tıbbi Bilişim Anabilim dalı Doktora programı çerçevesinde yürütülmüş olan “Meta Analizinde Heterojenliğin Saptanmasında Kullanılan Yöntemlerin Simülasyon Tekniği İle Karşılaştırılması” başlıklı çalışma, Jürimiz tarafından doktora tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 19.01.2011

Doç. Dr. E. Arzu KANIK

Mersin Üniversitesi Tıp Fakültesi

Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı

Jüri Başkanı

Prof. Dr. Nazan ALPARSLAN

Çukurova Üniversitesi Tıp Fakültesi

Biyoistatistik ve Tıbbi Bilişim AD.

Jüri Üyesi

Prof. Dr. Yusuf ÇELİK

Dicle Üniversitesi Tıp Fakültesi

Biyoistatistik ve Tıbbi Bilişim AD.

Jüri Üyesi

Doç. Dr. Bahar TAŞDELEN

Mersin Üniversitesi Tıp Fakültesi

Biyoistatistik ve Tıbbi Bilişim AD.

Jüri Üyesi

Yrd. Doç. Dr. Seva ÖNER

Mersin Üniversitesi Tıp Fakültesi

Halk Sağlığı AD.

Jüri Üyesi

Bu tez, Enstitü Yönetim Kurulunun 10/02/2011 tarih ve 2011/53 sayılı kararı ile kabul edilmiştir.

Prof. Dr. Ülkü CÖMELEK OĞLU



TEŐEKKÜR

Mersin Üniversitesi Saęlık Bilimleri Enstitüsünde tamamlamış olduęum “Meta Analizinde Heterojenlięin Saptanmasında Kullanılan Yöntemlerin Simülasyon Teknięi ile Karşılaştırılması” başlıklı doktora tezimi hazırlarken çalışmamın her aşamasında yakın ilgisi ve sonsuz sabrı ile bilimsel desteęini esirgemeyen Anabilim Dalı Başkanı ve danışman hocam, Sayın Doç. Dr. E.Arzu Kanık’a teşekkürlerimi sunarım. Ayrıca, doktora eğitimim boyunca deęerli bilgilerinden yararlandığım hocam Doç. Dr. Bahar TAŐDELEN’e teşekkür ederim.

Tez dönemi boyunca göstermiş olduęu anlayış ve sabırdan dolayı sevgili eşim Yrd. Doç. Dr. Eyüp ERDOĒAN’a teşekkür ederim.

Semra ERDOĒAN

Mersin, Ocak 2011

İÇİNDEKİLER

Kabul ve onay	i
TEŞEKKÜR.....	ii
İÇİNDEKİLER	iii
ŞEKİLLER DİZİNİ	v
ÇİZELGELER DİZİNİ	vii
ÖZET	viii
ABSTRACT.....	ix
1. GİRİŞ	1
2. GENEL BİLGİLER	5
2.1. Meta analizinde kullanılan kavramlar.....	5
2.1.1. Etki Büyüklüğü	5
2.1.1.1. Binary verinin özet istatistikleri (Sonuç değişkeni oran ise).....	5
2.1.1.1.1. Nisbi Risk (RR, Risk oranı, Risk hızı).....	6
2.1.1.1.2. Odds Oranı (OR, oranların oranı).....	7
2.1.1.1.3. Risk farkı (Risk difference (RD), Absolute risk reduction (ARR), Atfedilen Risk (AR))	9
2.1.1.1.4. Tedaviye gerekli sayı (NNT, The number needed to treat)	10
2.1.1.2. Sürekli yapıda kullanılan özet istatistikler.....	11
2.1.1.2.1. İki değişken arasındaki ilişkinin derecesi	11
2.1.1.2.2. İki ortalama arasındaki farklılığın derecesi	13
2.1.1.2.2.1. Ortalamalar arası fark (MD)	13
2.1.1.2.2.2. Standardize edilmiş ortalamalar arası fark.....	14
2.1.1.2.2.2.1. Cohen'in d istatistiği	14
2.1.1.2.2.2.2. Glass'ın Delta istatistiği	16
2.1.1.2.2.3. Hedges'in g istatistiği	17
2.1.1.2.2.4. Cevap oranı yöntemi (Response Ratio, Ratio of Means (RoM) Method).....	19
2.1.1.2.2.5. Ortak dilde etki büyüklüğü (CLES, Common Language Effect Size, CL) istatistiği	21
2.1.1.3. Etki Büyüklükleri arasındaki dönüşümler	22
2.1.1.4. Anlamlılık testinden yararlanarak etki büyüklüklerinin hesaplanması..	24
2.1.2. Ağırlık	25
2.1.3. Çalışmaların birleştirilmesinde kullanılan istatistiksel model seçimi.....	26
2.1.3.1. Sabit etki modeli (Fixed effect model)	26
2.1.3.2. Rasgele (Random) etki modeli	31
2.1.3.3. Sabit Etki Modeli ile Rasgele Etki Modelinin Karşılaştırılması.....	35
2.1.4. Meta Analizinde Heterojenlik.....	38
2.1.4.1. Heterojenliğin Testi	40
2.1.4.1.1. Cochran'ın Q istatistiği	40
2.1.4.2. Heterojenlik Ölçümleri	43
2.1.4.2.1. H istatistiği	43
2.1.4.2.2. τ^2 istatistiği.....	45
2.1.4.2.3. R istatistiği	47
2.1.4.2.4. I^2 İndeksi	48
2.1.4.3. Grafik gösterimleri.....	51

2.1.4.3.1. Forest grafik.....	51
2.1.4.3.2. Galbraith Radial grafik	52
2.1.4.3.3. L'abbe grafik.....	53
2.1.4.3.4. Funnel grafik.....	54
2.1.4.3.5. Normal quantile grafik.....	55
2.1.4.3.6. Standardize edilmiş artık histogramı	55
2.1.4.4. Meta Analizindeki Heterojenliğin Nedenleri.....	56
2.1.4.5. Meta Analizindeki Heterojenliğin Açıklanması	57
3. GEREÇ VE YÖNTEM.....	59
3.1. Meta Analizine Dahil Edilen Çalışmaların Homojen Olduğu Durumlar	59
3.1.1. Simülasyon Çalışması.....	59
3.2. Meta Analizine Dahil Edilen Çalışmaların Heterojen Olduğu Durumlar.....	60
3.2.1. Simülasyon Çalışması.....	60
4. BULGULAR.....	61
4.1. Meta Analizine Dahil Edilen Çalışmaların Homojen Olduğu Durumlarda Heterojenlik Testine ve Heterojenlik Ölçümlerine Ait Bulgular.....	61
4.2. Meta Analizine Dahil Edilen Çalışmaların Heterojen Olduğu Durumlarda Heterojenlik Testine ve Heterojenlik Ölçümlerine Ait Bulgular.....	70
5. TARTIŞMA.....	80
5.1. Heterojenlik Testi ve Heterojenlik Ölçümlerinin Meta Analizine Dahil Edilen Çalışma Sayısından ve Örneklem Büyüklüğünden Etkilenme Durumları	80
5.2. Heterojenlik Testine Ait Anlamlılık Seviyesi ile Heterojenlik Ölçümlerine Ait Kesim Noktalarının Belirlenmesi	90
6. SONUÇLAR VE ÖNERİLER.....	104
7. KAYNAKLAR	106
ÖZGEÇMİŞ.....	110

ŞEKİLLER DİZİNİ

Şekil 2.1. Sabit etki modeli-gerçek etki büyüklükleri	27
Şekil 2.2. Sabit etki modeli- gerçek etki büyüklüğü ve örnekleme hatası	28
Şekil 2.3. Sabit etki modeli- örnekleme hatasının dağılımı	29
Şekil 2.4. Rasgele etki modeli- gerçek etki büyüklüğünün dağılımı	32
Şekil 2.5. Rasgele etki modeli- tek bir çalışmadaki gerçek ve gözlenen etkiler	32
Şekil 2.6. Küçük örneklem büyüklüklerine sahip çalışmaların istatistiksel modellerdeki değişimi	37
Şekil 2.7. Büyük örneklem büyüklüklerine sahip çalışmaların istatistiksel modellerdeki değişimi	38
Şekil 2.8. Forest grafik	52
Şekil 2.9. Galbrait radial grafik	53
Şekil 2.10. L'abbe grafik	53
Şekil 2.11. Funnel grafik	54
Şekil 2.12. Normal quantile grafik	55
Şekil 2.13. Standardize edilmiş artık histogramı	56
Şekil 4.1. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen Cochran Q.....	67
Şekil 4.2. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen I^2	67
Şekil 4.3. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen H	68
Şekil 4.4. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen H_M	68
Şekil 4.5. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen τ^2	69
Şekil 4.6. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen R	69
Şekil 4.7. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen Cochran Q.....	77
Şekil 4.8. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen I^2	77
Şekil 4.9. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen H	78
Şekil 4.10. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen H_M	78
Şekil 4.11. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen τ^2	79
Şekil 4.12. Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen R	79
Şekil 5.1. Cochran Q'nun, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu	84
Şekil 5.2. I^2 test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu	85
Şekil 5.3. H test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu	86
Şekil 5.4. H_M test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu	87
Şekil 5.5. R test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu	88

Şekil 5.6. τ^2 test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu	89
Şekil 5.7. I^2 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	91
Şekil 5.8. H^2 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	92
Şekil 5.9. H_M 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	92
Şekil 5.10. R^2 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	93
Şekil 5.11. τ^2 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	93
Şekil 5.12. I^2 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	95
Şekil 5.13. H^2 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	95
Şekil 5.14. H_M 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	96
Şekil 5.15. R^2 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	96
Şekil 5.16. τ^2 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	97
Şekil 5.17. I^2 'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	98
Şekil 5.18. H^2 'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	99
Şekil 5.19. H_M 'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	99
Şekil 5.20. R^2 'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	100
Şekil 5.21. τ^2 'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması	100
Şekil 5.22: Farklı anlamlılık seviyelerinde ve çalışma sayılarında heterojenlik testinin gücü	103
Şekil 5.23: Farklı anlamlılık seviyelerinde ve örneklem büyüklüklerinde heterojenlik testinin gücü ..	103

ÇİZELGELER DİZİNİ

Çizelge 2.1. Etki büyüklüklerinin hesaplanmasında kullanılacak olan simgesel gösterimler.....	6
Çizelge 2.2. Gerçek ve gözlenen etki büyüklüklerinin gösterimleri	26
Çizelge 4.1. k=3 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları	63
Çizelge 4.2. k=5 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları	64
Çizelge 4.3. k=10 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları	65
Çizelge 4.4. k=20 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları	66
Çizelge 4.5. k=3 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları	73
Çizelge 4.6. k=5 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları	74
Çizelge 4.7. k=10 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları	75
Çizelge 4.8. k=20 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları	76
Çizelge 5.1. Cochran Q testi ve heterojenlik ölçümlerinin k ve N'ye göre korelasyon katsayıları ve p değerleri	81
Çizelge 5.2. Cochran Q'ya göre 0.05 anlamlılık seviyesinde heterojenlik ölçümlerine ait kesim noktaları	91
Çizelge 5.3. Cochran Q'ya göre 0.10 anlamlılık seviyesinde heterojenlik ölçümlerine ait kesim noktaları	94
Çizelge 5.4. Gerçek durumda heterojenlik testine ait anlamlılık seviyesi ve heterojenlik ölçümleri için kesim noktaları.....	98
Çizelge 5.5. 1000 denemede elde edilen Tip 1 hata yüzdeleri.....	101
Çizelge 5.6. 1000 denemede elde edilen Power yüzdeleri.....	101-102

ÖZET

Meta Analizinde Heterojenliğin Saptanmasında Kullanılan Yöntemlerin Simülasyon Tekniği ile Karşılaştırılması

Homojenlik analizi, etki büyüklüklerinin bir çalışmadan diğer bir çalışmaya nasıl değiştiğini gösteren bir ölçüttür. Meta analizi çalışmalarında heterojenliği test edebilmek amacıyla sadece Cochran Q test istatistiği ile değerlendirme yapılmasının yeterli olmadığı, aynı zamanda heterojenliğin etkisinin de heterojenlik ölçümleri ile tanımlanması tavsiye edilmektedir. Ancak, Cochran Q test istatistiğinin gücü, meta analizine dahil edilen çalışma sayısından etkilendiğinden, heterojenliğin olup olmamasına da klinik olarak önemli olup olmamasına da karar vermek zorlaşmaktadır. Cochran Q test istatistiğinin bu probleminden dolayı, anlamlılık seviyesi için 0.05 yerine 0.10 alınması tavsiye edilmektedir.

Yapılan simülasyon çalışması, meta analizinde kullanılan heterojenlik testi ve ölçümlerinin çalışma sayısından ve örneklem büyüklüğünden etkilenme durumları ortaya konmaya çalışılmış, heterojenlik testine ait anlamlılık seviyesi ve heterojenlik ölçümleri için uygun kesim değerleri hesaplanmıştır. Elde edilen bulgularının sonucuna göre, meta analizi çalışmalarında Cochran Q testine ait anlamlılık seviyesi olarak 0.05 veya 0.10 alınmıyorsa heterojenlik ölçümü olarak, gerçek homojen ve heterojenleri en iyi ayırt edebilen heterojenlik ölçümü olarak τ^2 test istatistiğinin kullanılması önerilmektedir. Bunun yanı sıra, Cochran Q test istatistiği için anlamlılık seviyesi olarak 0.30 alındığında tüm heterojenlik ölçümlerinin gerçek homojen ve heterojenleri çok iyi ayırt ettiği ortaya konulmuştur. Heterojenliğin test edilmesinde, heterojenlik ölçümlerinden I^2 test istatistiği, % 15.66'nın üzerinde bir değer aldığındaki, H ve R için 1.15'in üzerinde olması durumunda, H_M için 0.62'nin üzerinde olması, çalışmalar arası varyansın da (τ^2) 0.13'ün üzerinde olması durumlarında da meta analizine dahil edilen çalışmaların heterojen olduğu sonucuna varılabilmektedir.

Anahtar Kelimeler: Meta-Analizi, Heterojenlik, Heterojenlik ölçümleri, Cochran Q, I^2 index.

ABSTRACT

Methods used for the detection of heterogeneity in Meta-Analysis comparison with simulation technique

Homogeneity analysis, effect sizes are criterion that shows how a change in study to another study. Cochran Q test statistic is just not enough of making the assessment in Meta-analysis studies in order to test the heterogeneity, at the same time, the heterogeneity measures are recommend definition the effect of heterogeneity. However, the power of the Cochran Q test statistic, the numbers of studies are included in Meta analyses have been affected, in this case, it would be difficult to decide whether or not heterogeneity and whether or not clinical significance. Level of significance is recommended to be taken as 0.10 instead of 0.05 due to this property of Cochran Q heterogeneity test.

The simulation study, the heterogeneity test and heterogeneity measurement to put forward affected the number of study and sample size, calculated cut-off values for level of significance belonging to heterogeneity test (p) and for heterogeneity measurements. According to simulation results obtained from this study, Cochran Q test significance level of meta-analysis studies are being received as a 0.05 or 0.10 as a measure of heterogeneity, τ^2 test statistic used to recommend as a measure of heterogeneity the best that can distinguish between real homogeneous and heterogeneous. In addition, the Cochran Q test statistic if level of significance of Cochran Q value is above 0.30 for all measures of heterogeneity could distinguish very well put forward the real homogeneous and heterogeneous. Additionally, as the result of simulation obtained from this study, we can conclude that it is heterogeneous regardless of Cochran Q value if I^2 that is one of heterogeneity measurements is above 15.56. Besides, studies included in the meta-analyses are concluded to be heterogeneous if the other heterogeneity measurements H and R are above 1.15, H_M is above 0.62, variance between studies (τ^2) is above 0.13.

Key Words: Meta-Analysis, Heterogeneity, Measures of heterogeneity, Cochran Q statistics, I^2 index.

1. GİRİŞ

Bilimsel arařtırmalarda gvenilir sonulara ulařabilmek iin, alıřmanın iyi bir biimde planlanması, verilerin saėlıklı bir Őekilde toplanması, deėerlendirmede kullanılacak olan istatistiksel yntemlerin doėru seilmesi ve yorumlanması gerekmektedir.

Arařtırmacı, ilgilendiėi konuda yaptıėı arařtırma sonuları ile o konuda daha nceden yapılmıř olan alıřmaları inceleyip sonuları karřılařtırdıėında pek ok arařtırmanın sonularının birbirinden farklı sonuların olduėu durumla karřı karřıya kalmaktadır. Doėal olarak bu durum, alıřma yapacak olan kiřinin belirlediėi ama doėrultusunda ilerlemesini ve karar vermesini zorlařtıracaktır. Arařtırmacı byle bir durumda farklı olan bu sonulardan neye gre ve nasıl karar verecektir? (1).

alıřmalar arası ortaya ıkan bu farklılıkların eleřtirel deėerlendirilmesi, farklılıkların nereden kaynaklandıėının incelenmesi tıbbi bir bilginin kazanılmasında vazgeilmez bir unsurdur. Birbiriyle farklılık gsteren sonularda genel bir sonuca ulařabilmek iin yeni bir ynteme gereksinim duyulmuř ve meta analizi adı verilen istatistiksel bir yntem geliřtirilmiřtir. Farklı alıřmalardan elde edilen sonuların bir araya getirilmesi ile yapılan meta analizi, tek bir alıřmanın verdiėi sonulardan ok daha gvenilir ve etkin sonular vermektedir (2, 3).

Meta analizi, aynı konuda farklı yer, zaman ve merkezlerde yapılmıř olan arařtırma sonularını niteliksel ve niceliksel olarak birleřtirmeye ve o konuda genel bir sonuca ulařmaya yardımcı olan istatistiksel bir yntemdir (4).

Meta analizi her alanda olduėu gibi rasgele kontroll denemeler, deneysel alıřmalar, gzlemsel alıřmalar (kohort, vaka-kontrol alıřmaları), tanı testlerini deėerlendiren alıřmalar ve epidemiyolojik alıřmalar gibi tıbbi alanda yaygın olarak kullanılan tm arařtırma trlerinde de kullanılmaktadır. oėu meta analizi nedensel iliřkileri aıklamayı amalayan alıřmalarla ilgilenmesine raėmen, nedensel olmayan iliřkilerin alıřmalarına, tanımlayıcı arařtırmalara, tarama alıřmalarına, teřhis yntemlerinin geerlilik ve gvenilirlik alıřmalarına, maliyet-etkinlik alıřmalarına da uygulanabilmektedir (5).

Genel olarak meta analizinin bařlıca amaları ařaėıdaki gibi sıralanmaktadır (2);

- ✓ Bilimsel literatürde ortaya çıkan tutarsızlıkları değerlendirmek ve nedenlerini incelemek
- ✓ Küçük örneklerle yürütülmüş çalışmaları birleştirerek toplam örneklem büyüklüğünü artırarak parametre tahminlerinin kesinliğini ve gücünü artırmak
- ✓ Çalışmalar arasında ortaya çıkan heterojenliğin doğru kaynaklarını bulmak
- ✓ Bireysel çalışmalarda düşünülmeyen ancak etkisi olduğu varsayılan farklı alt gruplarda tedavi etkinliğini ve değişimini incelemek
- ✓ İleride yapılacak olan araştırmalara yardımcı olabilmek
- ✓ Elde edilen sonuçlara göre sonraki çalışmalara ışık tutmak ve yeni araştırma konuları ortaya atmak

Meta analizinin birçok avantajının yanında bir takım dezavantajları da vardır. Bunlar, meta analizinin araştırılan konu için büyük bir titizlik ve dikkat gerektirmesi, literatür taramasının çok zaman alıcı olması, bu analizin en önemli özelliği olan araştırılan konu üzerinde fikir birliğine ulaşılamamış, herhangi bir uzlaşma noktası bulunamamış çalışmaları bir araya getirmek için uygun çalışmalara ulaşılması ve bunların analize dahil edilmesi gibi sıralanabilir (6).

Pek çok soruna yönelik çözüm getirmeyi amaçlayan meta analizi bazı araştırmacılar tarafından yanlış anlaşılmaktadır. Tüm çalışmaların bir havuzda biriktirilerek hepsinden genel bir sonuca gidilmesi düşüncesi en sık rastlanılan hatadır. Aslında meta analize, tüm bilimsel araştırmalarda olduğu gibi, öncelikle ilgilenilen konuya ilişkin bir hipotez kurularak başlanır (sorunun tanımlanması). Ardından araştırmanın amacı ve hedefleri belirlenir. Araştırılan konuya ilişkin literatür taraması yapılır. Yayınlanmış veya yayınlanmamış literatür ve tamamlanmamış araştırma raporlarının hepsi literatür çalışmasının içine girmektedir. Literatür taraması sonucunda bulunan tüm çalışmalar meta analize dahil edildiğinde hatalı ve yanlış sonuçlar elde edilebilmektedir. Bunu önleyebilmek için araştırmacı, çalışmaya alınacak olan araştırmaları, çalışmanın başında belirlediği kabul ve red kriterlerini dikkate alarak belirlemelidir. Seçim kriterlerine karar verildikten sonra, konuyla ilgili en az iki uzmanın çalışmaları okuyup, sonuçlar hakkında fikir birliğine varıldıktan sonra çalışmaların meta analize dahil edilmesi gerekmektedir. Bu işlem meta analizinin geçerliği ve güvenilirliği açısından oldukça önemlidir. Ardından araştırmacı, tüm çalışmalardan ilgilenilen bilgilere yönelik parametre ve değişkenleri tablo halinde

sunmalıdır. Ve son olarak amaca yönelik olarak uygun meta analizi yöntemi kullanılarak sonuçlar yeniden gözden geçirilip, yorumlanmalı ve rapor haline getirilmelidir (7). Meta analizi çalışmalarının kalitesini artırmak amacıyla, Prisma 2009 adı verilen bir kontrol listesi yayınlanmıştır ve bir meta analizi çalışmasının yayınlanabilmesi için bu kontrol listesine uyulması gerekmektedir (8).

1950’li yıllardan itibaren, önceleri farklı yöntemlerle yapılan araştırmaların birleştirilmesi ve ardından benzer yöntemlerle yapılmış araştırmaların bir araya getirilmesi ile giderek sağlamlaşan bir teorik temel üzerinde hızlı bir gelişme gösteren meta analizi, son on yılda “kanıta dayalı tıp uygulamalarında” en güvenilir kanıt haline gelmiştir. İstatistik yöntemlerinin olgunlaşması, ilgili klinik araştırmaların büyük bölümünün analize dâhil edilebilmesi ve internet olanaklarının artması da meta analizinin güvenilirliğini büyük ölçüde artırmıştır (9).

Sistematik derleme ve meta analizi bir derleme türü olup karıştırılmaktadır. Sistematik derleme, ilgilenilen konuda yayınlanmış olan çalışmaların kapsamlı bir biçimde taranıp, çeşitli kabul ve red kriterleri kullanılarak, hangi çalışmaların derlemeye gireceğini belirlenip sentezlenmesinden oluşmaktadır. Belirli bir konu hakkında verilerin sistematik olarak toplanması, her zaman istenilen ve mümkün olan bir durumdur. Fakat farklı çalışmalarda yer alan verilerin istatistiksel yöntemler kullanılarak birleştirilmesi her zaman mümkün olamamaktadır. Buradan yola çıkarak, tüm sistematik derlemelerin meta analizi olmadığı söylenebilmektedir (10).

1992 yılında İngiliz Hükümeti tarafından mükemmel klinik uygulamaları belirlemek, klinik araştırmalara ilişkin epidemiyolojik veriyi derleyip araştırmacıların kullanımına sunmak amacıyla bir merkez kurulmuştur. Merkeze klinik çalışmalar yoluyla en doğru ve mükemmel uygulamaların (tedavilerin) belirlenebileceği düşüncesini savunan Epidemiyolog Archie Cochrane’nın adı verilmiştir. Cochrane Merkezi, internetin de yaygınlaşması ile çok geniş veritabanlarında klinik araştırmaların meta analizini yapabilecek, metodolojik standartları belirleyebilecek ve geniş bir tartışma ortamı oluşturacak bir platform haline dönüştürülmüştür (9). Meta analizi, 1989 yılından itibaren MEDLINE’da “Konu Başlığı (Subject heading)”, 1992’den itibaren “Yayın türü (Publication Type)” olarak yer almaktadır.

Meta analizi yalnızca ortalama tedavi etkilerini (analitik meta analizi) tahmin etmek için değil, aynı zamanda heterojenliğin (exploratory meta analizi) olası kaynaklarını arařtırmak için de çok önemlidir (11).

Homojenlik analizi, etki büyüklüklerinin bir çalışmadan diđer bir çalışmaya nasıl deđiřtiđini gösteren bir ölçüttür. Bir meta analizinde sonuçların birleřtirilmesinden sonra bu birleřtirilmiř bulguları kullanmadan önce heterojenlik için istatistiksel testlerin ve grafiksel gösterimlerin incelenmesi gerekmektedir. Eđer homojenlik sađlanamamıřsa çalışmalar ve sonuçlar arası tutarsızlık varsa böyle bir durumda heterojenlik uygun istatistiksel yöntemlerle analiz yapılarak çalışmanın homojenliđi sađlanır. Aksi takdirde bilgiler istatistiksel açıdan güvenilir olmayabilir (6, 8).

Bu amaçla, meta analizine dahil edilen her bir çalışmanın sonuç deđiřkeninin binary yapıda olduđu, binary yapıda kullanılan etki büyüklüğü $\log(OR)$ deđerinin 1'e eřit olduđu durumlar göz önünde bulundurularak simülasyonlar iki bölümde ele alınmıřtır. İlk bölümde, meta analizine dahil edilen çalışmaların homojen olduđu, ikinci bölümde de meta analizine dahil edilen çalışmaların heterojen olduđu durumlara ait, farklı örneklem büyüklükleri ve tedavi etkinlik oranlarında veriler üretilmiř ve bu verilere ait Cochran Q heterojenlik testi, testin anlamlılık deđer ve heterojenlik ölçümleri hesaplatılmıřtır. Bu simülasyon sonuçlarından yararlanarak, heterojenlik testi ve heterojenlik ölçümlerinin, meta analizine dahil edilen çalışma sayısından ve örneklem büyüklüğünden, her iki gruba ait tedavi etkinlik oranları arasındaki farktan etkilenip etkilenmedikleri ortaya konmaya çalışılmıřtır. Ayrıca, heterojenlik testine ait anlamlılık seviyesi için ve heterojenlik ölçümleri için ROC analizleri yapılarak en uygun kesim noktaları elde edilmeye çalışılmıřtır.

2. GENEL BİLGİLER

2.1. Meta analizinde kullanılan kavramlar

2.1.1. Etki Büyüklüğü

Etki büyüklüğü, meta analizinin temel konusu olup, ilk olarak 1977 yılında Cohen tarafından geliştirilmiştir. Etki büyüklüğü (effect size) veya etki genişliği (tedavi etkisi, treatment effect) genellikle iki değişken arasındaki ilişki miktarı veya iki grup arasındaki farklılığın miktarı olarak kullanılan bir indeks değeridir. Meta analizinde çalışmaya dahil edilen her bir çalışma sayısal olarak birleştirilmeden önce çalışmaların her biri için sonuç değişkenine ait ölçümlerin (etki büyüklüklerinin) hesaplanması gerekmektedir (12).

Sayısız etki büyüklüğü türleri vardır. Etki büyüklükleri verinin biçimine göre değişiklik göstermektedir. Meta analizine dahil edilen her bir çalışmanın sonuç değişkeni binary yapıda ise tedavi ve kontrol grubunda olayın (hastalığın görülme sıklığı, ölüm olasılığı, vb.) görülme sıklıkları arasındaki fark ile ilgilenilmektedir. Sonuç değişkeni sürekli bir yapıda ise, iki grup ortalamaları arasındaki farklılığın derecesi ve iki bağımsız değişken arasındaki ilişkinin derecesi olarak iki alt başlık altında toplanmaktadır (12). Ya da kısaca iki grup karşılaştırılırken sonuç değişkenimizin oran veya ortalama oluşuna göre etki büyüklükleri hesaplanmaktadır.

2.1.1.1. Binary verinin özet istatistikleri (Sonuç değişkeni oran ise)

İki grup arasında risk veya odds karşılaştırılabilir. En sık kullanılan özet istatistikleri, nisbi risk (risk oranı, RR), odds oranı (OR), Risk farkı (RD, Risk difference) ve tedaviye gerekli sayı (NNT, the number needed to treat) olarak verilmektedir. Özet istatistiklerinin hesaplanmasında Çizelge 2.1'deki gibi çapraz tablolardan yararlanılmaktadır.

Çizelge 2.1:Etki büyüklüklerinin hesaplanmasında kullanılacak olan simgesel gösterimler

	İlgilenilen olay		Toplam
	+	-	
Tedavi	a	b	a+b
Kontrol	c	d	c+d
Toplam	a+c	b+d	N

2.1.1.1.1. Nisbi Risk (RR, Risk oranı, Risk hızı)

Nisbi risk sağlık alanında yaygın olarak kullanılan bir risk ölçütüdür. Tedavi grubunda ilgilenilen olayın (hastalık, ölüm, vb) görülme olasılığının, kontrol grubunda ilgilenilen olayın görülme olasılığına oranlanması ile hesaplanır. Kısaca, tedavi grubunun ve kontrol grubunun risklerinin oranıdır ve “risk oranı” olarak da adlandırılır. Çizelge 2.1’den yararlanılarak aşağıdaki gibi formüle edilmiştir.

$$RR = \frac{\left[\frac{a}{(a+b)} \right]}{\left[\frac{c}{(c+d)} \right]} \quad (2.1)$$

Tedavi grubunda ilgilenilen olayın görülme olasılığı, kontrol grubunda ilgilenilen olayın görülme olasılığına eşit ise RR değeri 1’e eşit olur ve bu durum herhangi bir tedaviyi alan kişiler ile almayan kişiler arasında risk faktörüyle karşılaşma olasılıklarının eşit olduğu anlamına gelir. Tedavi grubunda ilgilenilen olayın görülme olasılığı, kontrol grubunda ilgilenilen olayın görülme olasılığından daha fazla ise RR değeri 1’den büyük çıkacaktır. Bu durum risk artışı göstermektedir. Aynı şekilde tedavi grubunda ilgilenilen olayın görülme olasılığı, kontrol grubunda ilgilenilen olayın görülme olasılığından daha az ise RR değeri 1’den küçük çıkacaktır ki bu durum ise riskin azaldığını ifade etmektedir (13, 14). Güven aralığı RR oranı etrafında simetrik değildir. Hiçbir zaman sıfırdan küçük bir değer almaz, ama oldukça büyük değerler alabilir. Buna karşın, logaritmik RR oranının güven aralığı yaklaşık olarak simetriktir. Bu nedenle güven aralığı RR’nin logaritmik değerine göre hesaplanır.

Ln (RR) için % 95 güven aralığı aşağıdaki gibi hesaplanmaktadır:

$$\ln(RR) \pm 1.96 * \sqrt{Var(\ln(RR))} \quad (2.2)$$

Formülde $\ln(RR)$ 'ye ait varyans değeri,

$$Var(\ln(RR)) = \frac{1}{a} - \frac{1}{(a+b)} + \frac{1}{c} - \frac{1}{(c+d)} \quad (2.3)$$

şeklinde hesaplanır. RR için güven aralığı ise aşağıdaki gibi hesaplanır;

$$\exp(\ln(RR) \pm 1.96 * \sqrt{Var(\ln(RR))}) \quad (2.4)$$

Özet istatistiğinin anlamlı olabilmesi için güven aralığının 1'i içermemesi gerekmektedir (14-16).

2.1.1.1.2. Odds Oranı (OR, oranların oranı)

Odds oranı da nisbi risk gibi bir risk katsayısıdır. Tedavi grubunda ilgilenilen olayın meydana gelme odds'unun kontrol grubunda ilgilenilen olayın meydana gelme odds'una oranı olarak basitçe ifade edilir. Odds oranı, vaka-kontrol (case-control) çalışmalarında, kohort çalışmalarında veya klinik denemelerde etki büyüklüğünün bir ölçüsü olarak yaygın bir şekilde kullanılmakta olup aşağıdaki gibi hesaplanmaktadır.

$$OR = \frac{\frac{a}{(a+b)}}{\frac{b}{(a+b)}} = \frac{\frac{a}{b}}{\frac{c}{(c+d)}} = \frac{ad}{bc} \quad (2.5)$$

RR gibi OR'nin de güven aralığı simetrik değildir. Odds oranı herhangi bir pozitif değer alabilir fakat negatif değer alamaz. Bu nedenle RR'de olduğu gibi Odds oranının

güven aralığı hesaplanırken logaritmasından yararlanır. Buna göre odds oranının % 95 güven aralığı;

$$\exp\left(\ln(OR) \pm 1.96 * \sqrt{Var(\ln(OR))}\right) \quad (2.6)$$

şeklinde hesaplanır. Formüldeki Var (ln(OR)) ise;

$$Var(\ln(OR)) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} \quad (2.7)$$

ile verilmektedir (13-15).

Odds oranının 1'den küçük olması, risk faktörünün ya da etkenin hastalık riskini azalttığını, 1'den büyük olması hastalık riskini artırdığını, 1 olması ise risk faktörünün hastalık ile ilişkili olmadığını ifade etmektedir.

Odds oranı genellikle nisbi risk ile karıştırılmakta ve nisbi risk olarak yorumlanmaktadır. Eğer tedavi grubunda ilgilenilen olayın görülme olasılığı fazla ise hem OR hem de RR değeri 1'den büyük olacaktır, fakat OR değeri RR değerinden daha büyük olacaktır. Eğer tedavi grubunda ilgilenilen olayın görülme olasılığı küçük ise hem OR hem de RR değeri 1'den küçük olacaktır, fakat OR değeri RR değerinden daha küçük olacaktır. İlgilenilen olayın görülme olasılığı kontrol ve tedavi grubunda aynı ise böyle bir durumda hem OR hem de RR değeri 1'e eşit olacaktır (13, 16).

Odds oranı ile nisbi risk oranı arasındaki ilişki aşağıdaki gibidir.

$$RR = OR \frac{1 + \frac{n_c}{n_d}}{1 + \frac{n_a}{n_b}} \quad (2.8)$$

Denklemden, n_a tedavi grubundaki hastaların sayısını, n_b tedavi grubundaki sağlıklı bireylerin sayısını, n_c kontrol grubundaki hastaların sayısını ve n_d ise kontrol grubundaki sağlıklı bireylerin sayısını vermektedir. Bu durumda denklem aşağıdaki gibi ifade edilir (16).

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

2.1.1.1.3. Risk farkı (Risk difference (RD), Absolute risk reduction (ARR), Atfedilen Risk (AR))

Tedavi ve kontrol grubundan oluşan ve sonucu binary olan bir denemede, uygulanan tedavinin kontrole göre etkinliğini belirlemenin en kolay yolu, tedavi ve kontrol grubundaki ilgilenilen olayın meydana gelme riski arasındaki farkı hesaplamaktır. Atfedilen risk, mutlak risk farkı (absolute risk reduction), atfedilen göreceli risk (attributable relative risk), atfedilen risk azaltımı (attributable risk reduction) ve risk farkı (risk difference) olarak da adlandırılmaktadır. Risk farkı, kontrol grubundaki ilgilenilen olayın meydana gelme riskinden tedavi grubundaki ilgilenilen olayın meydana gelme riski çıkarılarak hesaplanır. İnsidansa dayalı olarak hesaplandığından kohort tipi ve tedavisel çalışmalarda geçerli bir ölçüttür (13, 17, 18).

$$ARR = RD = \frac{c}{c+d} - \frac{a}{a+b} \quad \text{veya} \quad ARR = RD = \left| \frac{a}{a+b} - \frac{c}{c+d} \right| \quad (2.10)$$

Risk farkı pozitif de olabilir, negatif de olabilir. Bu nedenle, logaritma dönüşümüne gerek yoktur. Buna göre risk farkı için % 95 güven aralığı aşağıda gibi hesaplanmaktadır.

$$ARR \pm 1.96 * \sqrt{Var(ARR)} \quad (2.11)$$

Eşitlikteki risk farkına ait varyans değeri ise aşağıdaki şekilde ifade edilir.

$$Var(ARR) = \frac{\left(\frac{a}{a+b}\right)\left(1 - \left(\frac{a}{a+b}\right)\right)}{(a+b)} + \frac{\left(\frac{c}{c+d}\right)\left(1 - \left(\frac{c}{c+d}\right)\right)}{(c+d)} \quad (2.12)$$

ARR anlamlı olabilmesi için % 95 güven aralığının 0'ı içermemesi gerekmektedir. ARR'yi ve güven aralığını hesaplamak ve yorumlamak oldukça kolaydır (14, 16, 17).

2.1.1.1.4. Tedaviye gerekli sayı (NNT, The number needed to treat)

Tedavi etkisini göstermenin diğer bir yolu da NNT'dir. Genellikle klinik denemelerde yeni tedavinin kontrole veya referans bir tedaviye göre etkinliğini belirlemede kullanılır. Etkinliği görebilmek için en az kaç hastanın ele alınması gerektiğini ifade eder. NNT, risk farkının tersidir ve aşağıdaki şekilde hesaplanır (16, 17):

$$NNT = \frac{1}{ARR} \quad (2.13)$$

$$NNT = \frac{1}{\left[\frac{c}{(c+d)} - \frac{a}{(a+b)} \right]} \quad (2.14)$$

NNT değerinin 1 olması tedavi edilen tüm hastalarda aynı klinik sonucun görülebileceğini ifade etmektedir. 2 olması, iki kişiden birinde aynı tedavi sonucunun görülebileceğini ve bu tedavinin oldukça etkili olduğunu göstermektedir.

Meta analizinde NNT'yi kullanarak çalışmalarını birleştirmek çok da yaygın bir yöntem değildir. NNT için % 95 güven aralığı da ARR'nin güven aralığının tersinin alınması ile hesaplanır. Eğer $ARR = 0.1$ ve buna ait % 95 güven aralığı $[0.05; 0.15]$ olarak hesaplanmışsa, buna göre $NNT = 1/0.1 = 10$ olacak ve güven aralığı da $[1/0.05; 1/0.15]$ ile $[6.67; 20]$ olarak hesaplanacaktır. Benzer bir şekilde $ARR = 0.1$ ve % 95 güven aralığı da $[-0.05; 0.25]$ olarak hesaplandığında NNT değeri 10 ve NNT'ye ait % 95 güven aralığı ise $[-20; 4]$ olarak hesaplanır. Bu güven aralığında iki problem vardır. Birincisi, NNT değeri yalnızca pozitif değer almalıdır. İkinci problem ise NNT değeri güven aralığı içerisinde yer almamaktadır. Yani güven aralığı 10'u içermemektedir. Bununla birlikte, sonuçlar yalnızca istatistiksel olarak anlamlı olduğunda güven aralığını vermek yeterli değildir. Böyle problemleri bir durumla karşılaşıldığında McQuay ve Moore (1997), NNT değerini, güven aralığını vermeden kullanmayı önermektedir (16, 19).

NNT değeri yalnızca pozitif değer alır, fakat negatif değer alıyorsa bu durumda yan etkiye gerekli sayı (NNH, number needed to harm) olarak adlandırılır ve negatif NNT (NNH) değeri tedavi yönteminin zararlı etkiye sahip olduğunu ifade etmektedir. Kısaca, NNH değeri ilaçların yan etkilerini karşılaştırmak için yararlı olmaktadır ve bu değer ne kadar yüksek olursa o ilacın yan etkisinin de o kadar az olduğunu ifade etmektedir (16, 17).

ARR değeri 0'a yaklaştığında, yeni tedavi yöntemi ile standart tedavi (kontrol) arasında hemen hemen bir farklılık olmadığı anlamına gelir. ARR değeri 0 değerini aldığı anda NNT değeri ∞ değerini alacaktır, böylece ikinci örnekteki [-20; 4] güven aralığını yorumlamak zor olacaktır. Böyle bir durumda Altman (1998), güven aralığını iki parçaya bölerek bir çözüm önerisi getirmiştir. Birincisi, NNT için [- ∞ , -20] (ya da NNH için [20, ∞]) değer aralığı ile [4; ∞] değer aralığı ikinci parça olacak şekilde iki parçaya bölünüyor ve sonra bu iki parça birleştirilerek güven aralığı belirleniyor. NNH için [20, ∞], NNT değerleri [4, ∞] olarak belirlenir (17).

2.1.1.2. Sürekli yapıda kullanılan özet istatistikler

Meta analizinde sonuç değişkenimiz sürekli bir yapıda olduğunda kullanılan etki büyüklükleri, iki grup ortalamaları arasındaki farklılığın derecesi ve iki bağımsız değişken arasındaki ilişkinin derecesi olarak iki alt başlık altında toplanabilir. İki ortalama arasındaki farklılığın derecesi olarak kullanılan etki büyüklükleri, ortalamalar arası fark (mean difference, difference in means, MD) ve standardize edilmiş ortalamalar arası fark (standardized mean difference, SMD) olarak adlandırılırken, sürekli iki değişken arasındaki ilişkinin derecesi olarak kullanılan etki büyüklüğü ise korelasyona dayalı etki büyüklüğü istatistiği olarak adlandırılmaktadır (3, 20, 21).

2.1.1.2.1. İki değişken arasındaki ilişkinin derecesi

İki sürekli değişken arasındaki korelasyon katsayısı (r) etki büyüklüğü istatistiği olarak ifade edilmekte olup aşağıdaki gibi hesaplanmaktadır.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.15)$$

Korelasyon katsayısının varyansı ise,

$$V_r = \frac{(1-r^2)^2}{n-1} \quad (2.16)$$

şeklinde hesaplanmaktadır. Meta analizi çalışmalarında r korelasyon katsayısı değeri doğrudan hesaba katılmaz. Korelasyon katsayısı Fisher'in Z ölçeğine dönüştürülerek meta analizinde birleştirilir. Ancak Fisher'in Z ölçeği korelasyon katsayısının anlamlılık testi olan Z ile karıştırılmamalıdır. Korelasyon değeri r, Fisher'in z'sine aşağıdaki formülle dönüştürülür:

$$Z = 0.5 \ln \left(\frac{1+r}{1-r} \right) \quad (2.17)$$

Z'nin varyansı;

$$V_Z = \frac{1}{(n-3)} \quad (2.18)$$

şeklinde standart hatası ise,

$$SE_Z = \sqrt{V_Z} \quad (2.19)$$

şeklinde hesaplanır. Fisher'in Z değeri $-\infty$ ile $+\infty$ arasında bir değer alır. Z'nin negatif değerleri ters yönlü bir ilişkiyi, pozitif değerleri aynı yönlü ilişkiyi, 0 değeri ise ilişkinin olmadığını göstermektedir. Fisher'in Z ölçeği kullanıldığında korelasyonun varyansı dikkate alınmaz. Daha doğrusu analizlerde Fisher'in Z ölçeği ile Z'nin varyansı kullanılmaktadır. Bu değerlerden yararlanarak aşağıdaki eşitlikle tekrar r değerine ulaşılabilir (21, 22).

$$r = \frac{e^{2Z} - 1}{e^{2Z} + 1} \quad (2.20)$$

2.1.1.2.2. İki ortalama arasındaki farklılığın derecesi

2.1.1.2.2.1. Ortalamalar arası fark (MD)

Tedavi ve kontrol grubu olmak üzere iki gruba ait populasyon ortalamaları sırasıyla μ_t ve μ_c , varyansları da σ_t^2 ve σ_c^2 ile gösterilmektedir. İki grup ortalaması arasındaki farklılık aşağıdaki şekilde formüle edilir:

$$MD = d = \mu_t - \mu_c \quad (2.21)$$

İki grup ortalaması arasındaki farklılığın varyansı ise;

$$Var(MD) = Var(d) = S^2_{pooled} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \quad (2.22)$$

şeklinde hesaplanır. Burada n_1 , tedavi grubundaki denek sayısı, n_2 ise kontrol grubundaki denek sayısını göstermektedir.

$$S_{pooled} = \sqrt{\frac{(n_1 - 1)S_t^2 + (n_c - 1)S_c^2}{(n_1 + n_c - 2)}} \quad (2.23)$$

Böyle bir durumda iki grup ortalaması arasındaki farklılığın varyansı;

$$Var(MD) = \sqrt{\frac{(n_1 - 1)S_t^2 + (n_c - 1)S_c^2}{(n_1 + n_c - 2)}} * \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \quad (2.24)$$

formülasyonuna dönüşür. Ortalamalar arası farka ait % 95 güven aralığı ise aşağıdaki şekilde hesaplanır (14, 21).

$$MD \pm 1.96 * \sqrt{Var(MD)} \quad (2.25)$$

2.1.1.2.2.2. Standardize edilmiş ortalamalar arası fark

2.1.1.2.2.2.1. Cohen'in d istatistiği

Cohen'in d istatistiği Cohen (1969) tarafından geliştirilmiştir. Meta analizi çalışmalarında etki büyüklüğünü hesaplamak için çok yaygın olarak kullanılan ölçümlerden biridir ve literatürde farklı hesaplama yöntemleri bulunmaktadır. Bunlardan bir tanesi aşağıdaki gibidir.

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{pooled}} \quad (2.26)$$

Denklem 2.26'da, \bar{X}_1 , tedavi grubuna ait, \bar{X}_2 , kontrol grubuna ait ortalamaları göstermektedir. σ_{pooled} , her iki grup için ortak standart sapma değeridir ve aşağıdaki gibi hesaplanır (Burada σ_{pooled} , σ_{within} olarak, S_{pooled} , $\sqrt{MS_{within}}$ olarak ifade edilir):

$$\sigma_{pooled} = \sqrt{\left(\frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1 + n_2} \right)} \quad (2.27)$$

$$\sigma_{pooled} = S_{pooled} \sqrt{\left(\frac{n_1 + n_2 - 2}{n_1 + n_2} \right)} \quad (2.28)$$

$$\sigma_{pooled} = \sqrt{\left(\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \right)} \sqrt{\left(\frac{n_1 + n_2 - 2}{n_1 + n_2} \right)} \quad (2.29)$$

$$\sigma_{pooled} = \sqrt{\left(\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2} \right)} \quad (2.30)$$

Böyle bir durumda Cohen'in d istatistiği ise,

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2}\right)}} \quad (2.31)$$

şekline dönüşür (3,20,23-27).

Cohen'in d istatistiğinin varyansı;

$$V_d = \left(\frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2 - 2)}\right) \left(\frac{n_1 + n_2}{n_1 + n_2 - 2}\right) \quad (2.32)$$

şeklinde % 95 güven aralığı ise,

$$\frac{w_i d_i}{w_i} \pm 1.96 \sqrt{\frac{1}{w_i}} \quad (2.33)$$

şeklinde hesaplanır. Güven aralığı formülünde yer alan w_i , ağırlığı ifade etmektedir ve aşağıdaki gibi elde edilir (15, 22, 28).

$$w_i = \frac{2n_1 n_2 (n_1 + n_2)}{2(n_1 + n_2)^2 + n_1 n_2 d} \quad (2.34)$$

Cohen'in d istatistiği bazı kaynaklarda ise aşağıdaki şekillerde formüle edilmektedir.

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\right)}} \quad (2.35)$$

$n_1 = n_2 = n$ ise böyle bir durumda formül aşağıdaki gibi olur (3, 20, 29, 30).

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{S_1^2 + S_2^2}{2}\right)}} \quad (2.36)$$

Ayrıca, MIX, Comprehensive Meta Analysis (CMA) ve MedCalc paket programlarında standardize edilmiş etki büyüklüğü olan Cohen'in d istatistiği 2.36'daki eşitliğe göre hesaplanmaktadır.

Cohen'in d'si t test sonuçlarından da hesaplanabilir. Her bir gruptaki denek sayıları ve t değeri biliniyorsa aşağıdaki gibi formüle edilir: (29, 31).

$$d = t \sqrt{\left(\frac{n_1 + n_2}{n_1 n_2}\right) \left(\frac{n_1 + n_2}{n_1 + n_2 - 2}\right)} \quad (2.37)$$

2.1.1.2.2.2. Glass'ın Delta istatistiği

1976 yılında Gene V. Glass tarafından geliştirilmiştir. Ortalamalar arasındaki farklılık, yalnızca kontrol grubuna ait standart sapma değerine bölünerek bulunur. Bu formül, Glass'ın deltası olarak bilinmektedir. S_2 , kontrol grubuna ait standart sapma değerini göstermektedir.

$$\Delta = \frac{\bar{X}_1 - \bar{X}_2}{S_2} \quad (2.38)$$

Glass'ın delta istatistiği, populasyon parametresinin bir tahmin edicisi olarak kullanılmak istenirse formül aşağıdaki gibi gösterilir:

$$\Delta = \frac{\mu_1 - \mu_2}{\sigma_2} \quad (2.39)$$

Denklem 2.39'daki μ_1 ve μ_2 sırasıyla tedavi ve kontrol grubuna ait ölçümlerin ortalamalarıdır. σ_2 ise kontrol grubunun populasyon standart sapmasını ifade etmektedir (3, 20, 27, 29, 32).

Glass'ın delta istatistiğinin varyansı aşağıdaki gibi gösterilmektedir ve formülasyondaki n_1 , tedavi grubuna ait örnek genişliğini, n_2 ise kontrol grubuna ait örnek genişliğini ifade etmektedir.

$$V_{\Delta} = \frac{(n_1 + n_2)}{n_1 n_2} + \frac{\Delta^2}{2(n_2 - 1)} \quad (2.40)$$

% 95 güven aralığı ise;

$$\frac{w_i \Delta_i}{w_i} \pm 1.96 \sqrt{\frac{1}{w_i}} \quad (2.41)$$

şeklinde hesaplanır (15, 22, 28).

2.1.1.2.2.3. Hedges'in g istatistiği

Glass'ın delta istatistiği formülünde, gruplar arasındaki farklılık sadece kontrol grubunun standart sapması ile standardize edilmektedir. Buna rağmen, ortalamalar arası farklılık her iki grubun varyanslarına bağlıdır. Bu nedenle, Glass'ın delta istatistiği, tedavi ve kontrol grubu arasındaki değişkenlikteki farklılıklardan çok az etkilenmektedir. Bu özellik, her bir grup içerisindeki değişkenlik farklılaştığında etki büyüklüğü tahmininde yanlışlık yaratabilir. Bu amaçla, Larry Hedges (1981, 1982), her iki grubun değişkenliğine dayanarak bir ölçüm için kontrol grubunun standart sapmasını tedavi grubunun standart sapması ile değiştirmeyi önermektedir. Buna göre Hedges'in g istatistiği formülü aşağıdaki gibi verilmektedir (3, 20, 22, 29).

$$g = \frac{\bar{X}_1 - \bar{X}_2}{S_{pooled}} \quad (2.42)$$

Hedges'in g istatistiği formülünde yer alan toplanmış standart sapma (S_{pooled}) değeri,

$$S_{pooled} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \quad (2.43)$$

şeklinde formüle edilir ve böylece Hedges'in g istatistiği formülü aşağıdaki şekle dönüşür.

$$g = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\right)}} \quad (2.44)$$

Hedges'in g istatistiğinin varyansı;

$$V_g = \frac{(n_1 + n_2)}{n_1 n_2} + \frac{g^2}{2(n_1 + n_2 - 2)} \quad (2.45)$$

şeklinde hesaplanmaktadır (22).

Hedges'in g istatistiği popülasyona ilişkin tahmini ise G ile gösterilir ve aşağıdaki gibi formüle edilir:

$$G = \frac{\mu_e - \mu_c}{\sigma} \quad (2.46)$$

Hedges'in g'si tedavi ve kontrol grubuna ait örneklem varyanslarının birbirinden farklı olması durumundan etkilenmez, yalnızca örnek genişlikleri küçük olduğunda pozitif yanlılığa sahip olma eğilimindedir. Yani, etki büyüklüğünü gerçek değerinin üstünde bir değer olarak hesaplar. Bu yanlılık için bir düzeltme terimi olarak Hedges ve Olkin (1985) tarafından küçük örnek genişliklerini hesaba katan yeni bir formülasyon önerilmiştir ($n > 10$, hatta $n = 5$ 'e eşit olacak kadar küçük olduğunda bile, bazı kaynaklarda da $n < 20$). Bu formülasyon g^* şeklinde verilebildiği gibi bazı kaynaklarda Hedges'in d istatistiği olarak da ifade edilmektedir (15, 22, 27, 29-32).

$$g^* = g \left[1 - \frac{3}{4(n_1 + n_2 - 2) - 1} \right] \quad (2.47)$$

$$g^* = g \left[1 - \frac{3}{4(n_1 + n_2) - 9} \right] \quad (2.48)$$

Bu istatistiğe ait varyans değeri;

$$V_d = V_{g^*} = \frac{n_1 + n_2}{n_1 n_2} + \frac{(g^*)^2}{2(n_1 + n_2)} \quad (2.49)$$

% 95 güven aralığı ise,

$$\frac{w_i g_i^*}{w_i} \pm 1.96 \sqrt{\frac{1}{w_i}} \quad (2.50)$$

şeklinde hesaplanır. Burada w_i değeri de aşağıdaki gibi formüle edilir (15).

$$w_i = \frac{2n_1 n_2 (n_1 + n_2)}{2(n_1 + n_2)^2 + n_1 n_2 (g_i^*)^2} \quad (2.51)$$

2.1.1.2.2.4. Cevap oranı yöntemi (Response ratio, Ratio of Means (RoM) Method)

Yukarıda bahsedilen üç farklı standardize edilmiş ortalamalar arası farka ait etki büyüklüğü istatistiklerine alternatif olarak Hedges ve ark. tarafından (1999) ortaya atılmış bir etki büyüklüğü ölçümüdür. Bazı kaynaklarda cevap oranı (response ratio) yöntemi olarak bazı kaynaklarda da ortalamaların oranı (ratio of means) yöntemi olarak adlandırılmaktadır ve bu etki büyüklüğü ölçümünde, tedavi ve kontrol grubu ortalamalarının farklılıkları yerine tedavi grubuna ait sonuç değişkeninin ortalama değerinin kontrol grubundaki sonuç değişkeninin ortalama değerine oranlanması

şeklinde hesaplanır. Bu yöntem için meta analizine dahil edilecek olan her bir çalışmaya ait cevap oranı değerleri hesaplanır ve sonra bu değerln logaritması alınır. Bu etki büyüklüğü ölçümü aşağıdaki gibi hesaplanmaktadır (21, 22, 33).

$$RoM = R = \frac{mean_{exp}}{mean_{control}} = \frac{\bar{x}_1}{\bar{x}_2} \quad (2.52)$$

$$\ln(RoM) = \ln(R) = \ln\left(\frac{\bar{x}_1}{\bar{x}_2}\right) = \ln(\bar{x}_1) - \ln(\bar{x}_2) \quad (2.53)$$

Cevap oranına ait varyans değeri ise aşağıdaki şekilde formüle edilir:

$$Var\left[\ln\left(\frac{\bar{x}_1}{\bar{x}_2}\right)\right] = Var[\ln(\bar{x}_1) - \ln(\bar{x}_2)] \quad (2.54)$$

$$= Var[\ln(\bar{x}_1)] + Var[\ln(\bar{x}_2)], \text{ (gruplar birbirinden bağımsız olduğunda)} \quad (2.55)$$

$$= \left(\frac{1}{\bar{x}_1}\right)^2 Var(\bar{x}_1) + \left(\frac{1}{\bar{x}_2}\right)^2 Var(\bar{x}_2) \quad (2.56)$$

$$Var(\bar{x}_i) = \frac{Var(x_i)}{n_i} = \frac{S_i^2}{n_i}, \text{ (x rd. için)} \quad (2.57)$$

$$= \frac{1}{n_1} \left(\frac{S_1}{\bar{x}_1}\right)^2 + \frac{1}{n_2} \left(\frac{S_2}{\bar{x}_2}\right)^2 \quad (2.58)$$

Cevap oranı yöntemine ait % 95 güven aralığı ise aşağıdaki şekilde hesaplanır (33).

$$\% 95 CI = \exp\left\{\left[\ln R\right] \pm 1.96 \sqrt{Var[\ln R]}\right\} \quad (2.59)$$

$$= \exp \left\{ \left[\ln \left(\frac{\bar{x}_1}{\bar{x}_2} \right) \right] \pm 1.96 \sqrt{\text{Var} \left[\ln \left(\frac{\bar{x}_1}{\bar{x}_2} \right) \right]} \right\} \quad (2.60)$$

Etki büyüklüğü istatistiklerinin hepsi $-\infty$ ile $+\infty$ arasında değer alır. Etki büyüklüğü istatistiğinin 0'a eşit olması tedavi ve kontrol grubu arasında etki bakımından farklılığın olmadığını, negatif bir değer alması kontrol grubunun tedavi grubuna nazaran daha etkili olduğunu, pozitif bir değer alması ise tedavi grubunun kontrol grubundan daha etkili olduğu anlamına gelmektedir (22).

2.1.1.2.2.5. Ortak dilde etki büyüklüğü (CLES, Common Language Effect Size, CL) istatistiği

İki bağımsız örneğin ortalamalarını karşılaştırırken etki büyüklüğünü tahmin etmede kullanılan başka bir yöntem ise ortak dilde etki büyüklüğü (CLES, Common Language Effect Size) istatistiği McGraw ve Wong (1992) tarafından geliştirilmiştir (34). Bu istatistik diğerlerinden daha kolay yorumlanmakta ve farklılığın büyüklüğü bir olasılık olarak ifade edilmektedir. Daha doğrusu CLES istatistiği bir olasılık tahminidir ve tedavi grubundan rasgele olarak seçilmiş bireye ait skorun, kontrol grubundan rasgele seçilmiş bireye ait skordan daha büyük olacağını ifade eden bir olasılıktır. Z skoru aşağıdaki gibi hesaplanmaktadır (29,34).

$$Z = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{(S_1^2 + S_2^2)}} \quad (2.61)$$

Buradan bulunan z değerinin standart normal dağılım tablosundan olasılık değeri belirlenir ve bu olasılık değeri 1'den küçük bir değerdir. Örneğin $Z=0.652$ olarak bulunduğu varsayalım. Buna karşılık gelen olasılık değeri ise $P(Z < 0.652) = 0.743$ olarak elde edilir ve bu değer % ile ifade edilir. % 74.3 olarak kolaylıkla yorumlanır. Tedavi grubundan rasgele seçilmiş bir bireyin, kontrol grubundan rasgele çekilmiş bireyin skorlarından % 74 kat daha büyük olduğu sonucuna varılmaktadır. Bir olasılık olarak etki büyüklüğünün bu dönüşümü, Cohen'in d etki büyüklüğü gibi diğer etki

büyükliklerine de kolaylıkla dönüştürülebilir. Standardize edilmiş etki büyüklüğü değerini $\sqrt{2}$ 'ye bölerek de bu istatistik için bir Z değeri hesaplanmış olur (29,34).

2.1.1.3. Etki Büyüklükleri arasındaki dönüşümler

Tek bir çalışma için etki büyüklüğünün hesaplanması hakkında unutulmaması gereken iki önemli nokta vardır. Bunlardan birincisi, etki büyüklüğü indekslerinin, ölçümün ortalama ve standart sapma değeri bilinmediğinde anlamlılık testlerinden hesaplanabilmesidir. Örneğin, bir çalışmada tedavi ile kontrol grubu karşılaştırıldığında, ortalama ve standart sapma değerleri verilmeyip sadece t testinin sonuçları verilir ise etki büyüklüğü istatistiği t değerinden hesaplanabilir. İkinci önemli nokta ise, etki büyüklüğü indekslerinin, diğer etki büyüklüğü indekslerine dönüştürülebilmesidir (3, 20).

Cohen d'den Hedges'in g istatistiğine dönüştürme;

$$g = d \sqrt{\left(\frac{df}{n_1 + n_2} \right)} \quad (2.62)$$

Cohen d'den r'ye dönüştürme;

$$r = \frac{d}{\sqrt{d^2 + \frac{(n_1 + n_2)^2}{n_1 n_2}}} \quad (2.63)$$

$n_1 = n_2$ ise,

$$r = \frac{d}{\sqrt{d^2 + 4}} \quad (2.64)$$

Hedges'in g istatistiğinden d'ye dönüştürme;

$$d = \frac{g}{\sqrt{df_{error}}} = g \sqrt{\frac{n_1 + n_2}{df}} \quad (2.65)$$

Hedges'in g istatistiğinden r'ye dönüştürme;

$$r = \sqrt{\frac{g^2 n_1 n_2}{g^2 n_1 n_2 + (n_1 + n_2) df}} \quad (2.66)$$

r'den Hedges'in g istatistiğine dönüştürme;

$$g = \frac{r}{\sqrt{(1-r^2)}} \sqrt{\frac{df (n_1 + n_2)}{n_1 n_2}} \quad (2.67)$$

r'den Cohen'in d'sine dönüştürme;

$$d = \frac{2r}{\sqrt{(1-r^2)}} \quad (2.68)$$

Glass Δ istatistiğinden d'ye dönüştürme;

$$d = \frac{\Delta}{\sqrt{df_{error}}} \quad (2.69)$$

Glass Δ istatistiğinden r'ye dönüştürme;

$$r = \sqrt{\frac{\Delta^2 n_1 n_2}{\Delta^2 n_1 n_2 + (n_1 + n_2) df}} \quad (2.70)$$

şeklinde yapılmaktadır (3).

Sürekli yapıda kullanılan etki büyüklükleri ile binary yapıda kullanılan etki büyüklükleri arasında da dönüşümler yapılabilir (21). Odds oranı istatistiğinden standardize edilmiş ortalamalar arası fark istatistiğine dönüştürme;

$$\text{Log}(OR) = d \frac{\pi}{\sqrt{3}} \quad (2.71)$$

şeklinde yapılmaktadır. Denklem 2.71’de π değeri matematiksel sabit olup 3.14159’a eşittir. Log (OR)’ye ait varyans ise aşağıdaki gibi hesaplanmaktadır.

$$V_{\text{Log}(OR)} = V_d \frac{\pi^2}{3} \quad (2.72)$$

Standardize edilmiş ortalamalar arası fark istatistiğinden log(OR) istatistiğine dönüştürme,

$$d = \text{Log}(OR) \frac{\sqrt{3}}{\pi} \quad (2.73)$$

şeklinde, d’ye ait varyans değeri ise Denklem 2.74’deki gibi formüle edilmektedir.

$$V_d = V_{\text{Log}(OR)} \frac{3}{\pi^2} \quad (2.74)$$

2.1.1.4. Anlamlılık testinden yararlanarak etki büyüklüklerinin hesaplanması

Cohen’in d istatistiğine dönüştürme;

$$d = t \frac{(n_1 + n_2)}{\sqrt{n_1 n_2} \sqrt{df}} \quad (2.75)$$

$n_1 = n_2$ ise,

$$d = \frac{2t}{\sqrt{df}} \quad (2.76)$$

Hedges'in g istatistiğine dönüştürme;

$$g = t \sqrt{\frac{(n_1 + n_2)}{n_1 n_2}} \quad (2.77)$$

$n_1 = n_2$ ise,

$$g = \frac{2t}{\sqrt{n_1 + n_2}} \quad (2.78)$$

r istatistiğine dönüştürme;

$$r = \sqrt{\frac{t^2}{t^2 + df}} \quad (2.79)$$

şeklinde yapılmaktadır (22-24, 35).

2.1.2. Ağırlık

Çoğunlukla, meta-analizine dahil edilen her çalışma, farklı örnek genişliklerine dayanmaktadır. Çok büyük örnek genişliklerine sahip çalışmaların tahminleri daha küçük örnek genişliğine sahip çalışmaların tahminlerinden daha iyi olması muhtemeldir. Analize katılan çalışmalardaki toplam denek sayısı, meta analizine esas alınan tüm çalışmalardaki toplam denek sayısına bölünerek her bir çalışmanın meta analizindeki ağırlığı hesaplanmaktadır. Etki büyüklükleri aynı zamanda çalışmanın kalitesi gibi diğer önemli göstergelerle hesaplanabilir (3).

2.1.3. Çalışmaların birleştirilmesinde kullanılan istatistiksel model seçimi

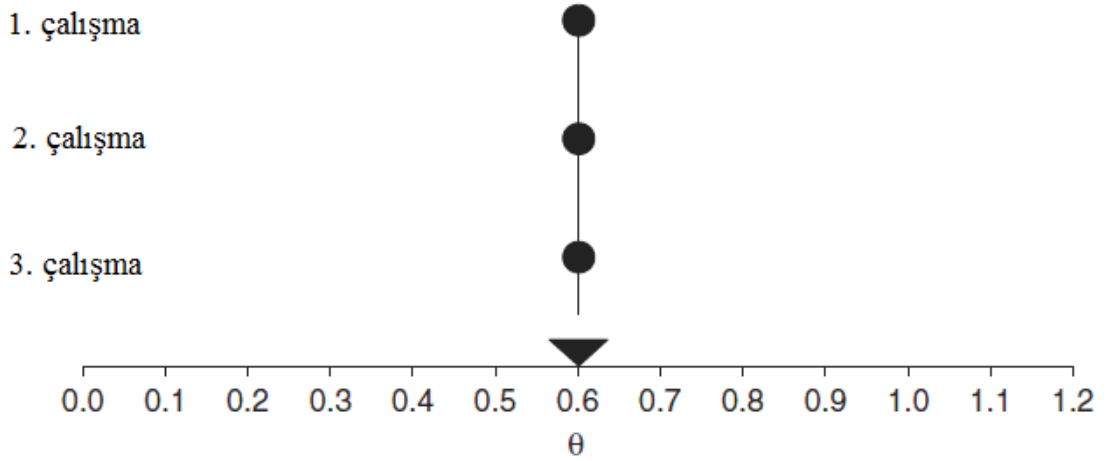
Çalışma sonuçlarının birleştirilmesinde istatistiksel model seçimi etkili olmaktadır. Meta analizleri, sabit etki modeli (fixed effect model) ve rasgele etki modeli (random effects model) olmak üzere iki istatistiksel modelden biri temeline dayanmaktadır. Model seçiminde gözlenen etki büyüklüğü ve gerçek etki büyüklüğü kavramlarının açıklanması gerekmektedir. Bir çalışmanın gerçek etki büyüklüğü popülasyondaki etki büyüklüğü, gözlenen etki büyüklüğü ise gerçekten gözlenen etki büyüklüğüdür. Eğer çalışma çok geniş bir örnekleme sınırlandırılmış bir çalışma ise böyle bir durumda örnekleme hatası sıfır olacağından dolayı gözlenen etki büyüklüğü ile gerçek etki büyüklüğü birbirine eşit olacaktır. Bir çalışmanın gözlenen ve gerçek etki büyüklükleri Çizelge 2.2’de verildiği gibi sembolize edilmektedir (21).

Çizelge 2.2: Gerçek ve gözlenen etki büyüklüklerinin gösterimleri

	Gerçek etki büyüklüğü (true effect size)	Gözlenen etki büyüklüğü (observed effect size)
Çalışma	●	■
Birleştirilmiş çalışmalarında	▼	◆

2.1.3.1. Sabit etki modeli (Fixed effect model)

Bu modeldeki temel varsayım, meta analizine alınacak olan her bir çalışmanın tamamen ortak bir etki büyüklüğüne sahip olmasıdır. Yani, etki büyüklüğünü etkileyen bütün faktörler meta analizine dahil edilen tüm çalışmalarda aynı olduğu için gerçek etki büyüklüğü bütün çalışmalarda sabittir. Gerçek etki büyüklüğü θ ile gösterilmektedir ve tüm çalışmalardaki gerçek etki büyüklüklerinin ortalamasına eşittir. Şekil 2.1’de, her bir çalışmaya ait etki büyüklükleri daire ile meta analizindeki etki büyüklüğü ise üçgen ile gösterilmektedir. Sabit etki modeli varsayımı altında, her bir çalışma için gerçek etki büyüklüğü ile birleştirilmiş çalışmalardaki gerçek etki büyüklükleri eşit olup bunlara ait daireler üçgenin üstünde yer almaktadır (21).

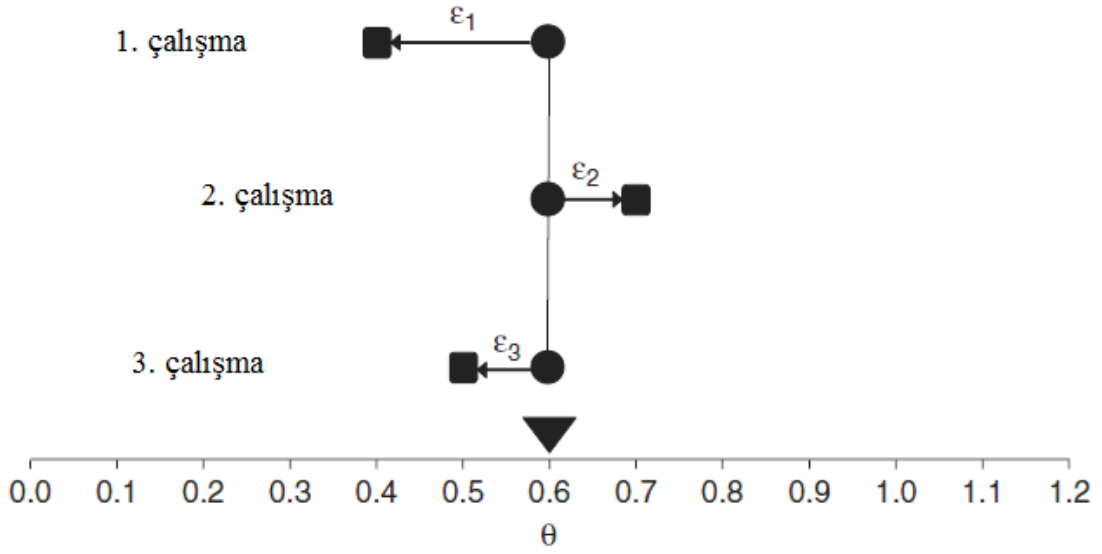


Şekil 2.1: Sabit etki modeli-gerçek etki büyüklükleri

Bütün çalışmalarda gerçek etki büyüklükleri aynı iken, her bir çalışmaya ait gözlenen etki büyüklükleri, rasgele hatadan dolayı çalışmadan çalışmaya farklılık göstermektedir. Eğer her bir çalışmadaki örnek genişliği çok büyük ise, böyle bir durumda örnekleme hatası sıfır olacaktır ve her bir çalışmaya ait gözlenen etki büyüklükleri ile gerçek etki büyüklükleri birbirine eşit olacaktır. Her bir çalışmadaki örnek genişliği çok büyük değilse, böyle bir durumda örnekleme hatası oluşacak ve çalışmalardaki gözlenen etki büyüklükleri gerçek etki büyüklükleri ile eşit olmayacaktır. Şekil 2.2’de her bir çalışmadaki gerçek etki büyüklüğü 0.60 iken gözlenen etki büyüklükleri her bir çalışmada farklılık göstermektedir (21,36,37).

Genellikle, her bir çalışma için gözlenen etki büyüklüğü (Y_i), her bir çalışmaya ait gerçek etki büyüklüğü (populasyon ortalaması) ile çalışmanın örnekleme hatasının toplamından oluşmaktadır ve aşağıdaki gibi ifade edilir (21):

$$Y_i = \mu + \varepsilon_i \quad (2.80)$$



Şekil 2.2: Sabit etki modeli- gerçek etki büyüklüğü ve örnekleme hatası

Şekil 2.2'ye göre, 1. çalışmada örnekleme hatası (ϵ_1) -0.20, gözlenen etki büyüklüğü (Y_1) ise;

$$Y_1 = 0.60 - 0.20 = 0.40$$

2. çalışmada örnekleme hatası (ϵ_2) 0.10, gözlenen etki büyüklüğü (Y_2) ise;

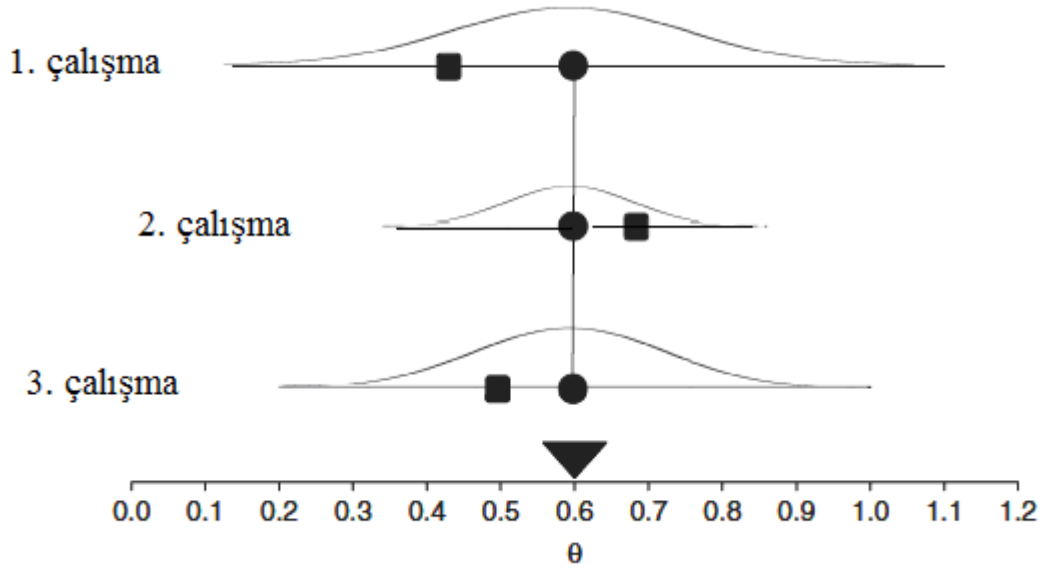
$$Y_2 = 0.60 + 0.10 = 0.70$$

3. çalışmada örnekleme hatası (ϵ_3) -0.10, gözlenen etki büyüklüğü (Y_3) ise;

$$Y_3 = 0.60 - 0.10 = 0.50$$

olarak hesaplanır (21,36,37). Herhangi bir çalışmada hata rasgele ise bu durumda hatanın örnekleme dağılımı tahmin edilmeye çalışılır. Şekil 2.3'de her bir çalışma için gerçek etki büyüklüğü etrafında dağılım eğrileri gösterilmektedir. Birinci çalışmada örnek genişliği küçük, varyans değeri büyük ve dolayısıyla gözlenen etki büyüklüğünün 0.2 ile 1.00 arasındaki herhangi bir noktaya düşmesi olasıdır. Bunun aksine ikinci

çalışmada örnek genişliği büyük, varyans değeri küçüktür ve gözlenen etki büyüklüğünün 0.4 ile 0.8 arasındaki dar bir mesafeye düşmesi olasıdır (21,36,37).



Şekil 2.3: Sabit etki modeli- örnekleme hatasının dağılımı

Meta analizi çalışmalarında gözlenen etki büyüklüğünden yararlanılarak populasyon etki büyüklüğü tahmin edilmeye çalışılır. Populasyon etki büyüklüğünün gerçek değerini tahmin etmek için ağırlıklı ortalama (M) hesaplanır ve bu değer aşağıdaki gibi formüle edilir (15, 21):

$$M = \frac{\sum_{i=1}^k w_i Y_i}{\sum_{i=1}^k w_i} \quad (2.81)$$

Burada Y_i , i. çalışmadaki gözlenen tedavi etkisinin değeridir (kullanılan etki büyüklüğü indeksi) ve $N(\theta, v_i)$ ile normal dağılıma sahiptir. k, meta analizine dahil edilen çalışma sayısı, w_i ise her bir çalışmanın ağırlığını vermektedir. Her bir çalışmanın ağırlığı, o çalışmaya ait varyansın (çalışma içi varyans) tersi olarak ifade edilmekte olup, $w_i = \frac{1}{v_{Y_i}}$ şeklinde hesaplanır. Böylece, birleştirilmiş çalışmalara ait

toplam etki büyüklüğünün varyansı, ağırlıkların toplamının tersine, standart hatası da varyansın kareköküne eşittir (15, 21).

$$v_M = \frac{1}{\sum_{i=1}^k w_i} \quad (2.82)$$

$$SE_M = \sqrt{v_M} \quad (2.83)$$

% 95 Güven aralığı ise Denklem 2.84'deki gibi verilmektedir.

$$M \pm 1.96 SE_M \quad (2.84)$$

Buna göre, sabit etki modelinde, etki büyüklükleri ve etki büyüklüklerine ait % 95 güven aralıkları aşağıdaki formülasyonlarla verilmektedir. OR ve % 95 güven aralığı;

$$OR = \exp\left[\frac{\sum_{i=1}^k w_i \ln(OR)}{\sum_{i=1}^k w_i}\right] \quad (2.85)$$

$$\% 95 \text{ Güven Aralığı} = \exp[OR \pm 1.96\sqrt{1/w_i}] \quad (2.86)$$

şeklinde hesaplanır. RR ve % 95 güven aralığı;

$$RR = \exp\left[\frac{\sum_{i=1}^k w_i \ln(RR)}{\sum_{i=1}^k w_i}\right] \quad (2.87)$$

$$\% 95 \text{ Güven Aralığı} = \exp[RR \pm 1.96\sqrt{1/w_i}] \quad (2.88)$$

şeklinde hesaplanır. ARR ve % 95 güven aralığı;

$$ARR = \frac{\sum_{i=1}^k w_i ARR}{\sum_{i=1}^k w_i} \quad (2.89)$$

$$\% 95 \text{ Güven Aralığı} = ARR \pm 1.96\sqrt{1/w_i} \quad (2.90)$$

şeklinde hesaplanır. NNT ve % 95 güven aralığı ise,

$$NNT = \frac{\sum_{i=1}^k w_i NNT}{\sum_{i=1}^k w_i} \quad (2.91)$$

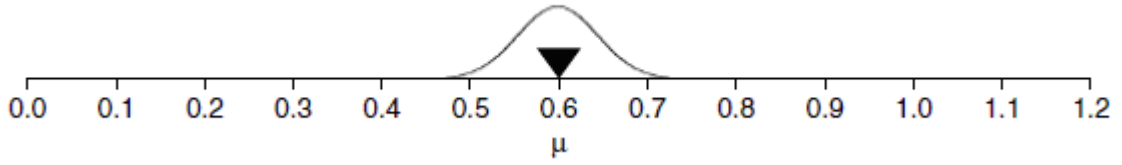
$$\% 95 \text{ Güven Aralığı} = NNT \pm 1.96\sqrt{1/w_i} \quad (2.92)$$

şeklinde hesaplanır (15, 21).

2.1.3.2. Rasgele (Random) etki modeli

Meta analizi çalışmalarında karar verilirken, meta analize dahil edilen çalışmaların etki büyüklüklerinin benzer olması varsayılır, fakat genellikle gerçek etki büyüklüğünün bütün çalışmalarda tamamen benzer olduğunu varsaymak için bir neden yoktur (15, 21).

Rasgele etki modelinin kullanıldığı meta analizi çalışmalarında gözlenen etki büyüklüğü (Y_i), gerçek etki büyüklüğü (θ_i) ve σ^2 ile normal dağılım gösterirken, buradaki θ_i 'de μ ortalama ve τ^2 varyans ile normal dağılım göstermektedir. Örneğin, Şekil 2.4'de bütün gerçek etki büyüklüklerinin ortalaması 0.60 fakat her bir çalışmadaki etki büyüklükleri değerleri ise normal dağılım eğrisi ile gösterildiğinde ortalama etrafında dağılmaktadır (21).

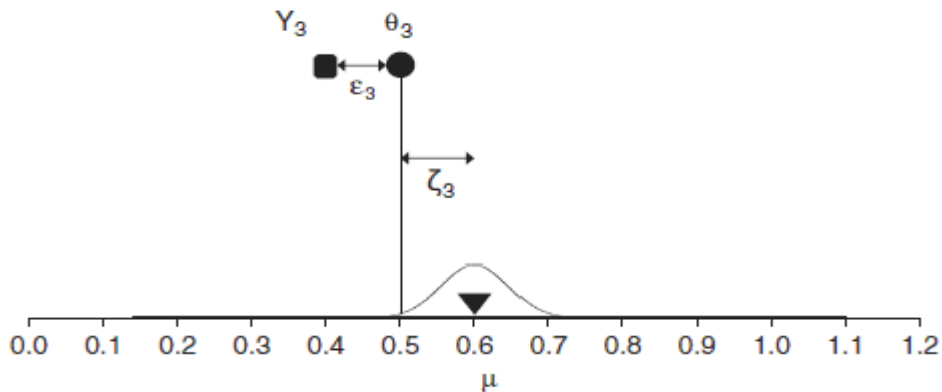


Şekil 2.4: Rasgele etki modeli- gerçek etki büyüklüğünün dağılımı

Sabit etki modelinde, her bir çalışma için gözlenen etki büyüklüğü, gerçek etki büyüklüğü ile örnekleme hatasının toplamından oluşmaktadır. Rasgele etki modelinde ise her bir çalışma için gözlenen etki büyüklüğü (Y_i), çalışmanın gerçek etki büyüklüğü (θ_i) ve populasyon ortalaması (μ) arasındaki değişim [$(\theta - \mu) = \zeta_i$] ile çalışmanın gerçek etki büyüklüğü ve gözlenen etki büyüklüğü arasındaki değişimin [$(Y_i - \theta) = \varepsilon_i$] toplamı olarak verilmekte olup aşağıdaki gibi formüle edilir (21).

$$Y_i = \mu + \zeta_i + \varepsilon_i \quad (2.93)$$

Şekil 2.5’de verilen çalışmanın gözlenen etkisi ile populasyon ortalaması arasındaki uzaklık iki belirgin kısımdan meydana gelmektedir ki bunlar etki büyüklüğündeki (ζ_i) ve örnekleme hatasındaki (ε_i) gerçek değişkenlerdir. Bu çalışmada, gerçek etki büyüklüğü (θ_3) 0.50, örnekleme hatası -0.10, gözlenen etki büyüklüğü 0.40, Y_3 ile populasyon ortalaması arasındaki uzaklık ise -0.20’dir (21).



Şekil 2.5: Rasgele etki modeli- tek bir çalışmadaki gerçek ve gözlenen etkiler

Her bir çalışmaya ait gözlenen etki büyüklüğünün populasyon ortalamasından uzaklığı çalışmalar arası gerçek etki büyüklüğünün standart sapmasına dayanmaktadır ve bu standart sapma değeri, çalışmalar arası değişkenlikle ilgili hata terimine ait standart sapma değeri olup τ (veya varyansı ise τ^2) ile gösterilir. Eğer her bir çalışma için gerçek etki büyüklüğü biliniyorsa ve bu etki büyüklüklerinin varyansı hesaplanabiliyorsa bu varyans çalışmalar arası varyanstır (τ^2) (21).

Sabit etki modelinde de rasgele etki modelinde de her bir çalışma varyansın tersi ile ağırlıklandırılır. Rasgele etki modelindeki farklılık, bu etki modelindeki çalışmanın varyansı, çalışma içi varyans ile çalışmalar arası varyansın (τ^2) toplamına eşit olmasıdır. Rasgele etki modelinde, her bir çalışmanın ağırlığı sabit etki modelindeki ile aynıdır sadece rasgele etki modelindeki formülasyonlar * işareti ile gösterilmektedir (21).

$$w_i^* = \frac{1}{v_{Y_i}^*} \quad (2.94)$$

Burada $v_{Y_i}^*$, i. çalışmaya ait çalışma içi varyans ile çalışmalar arası varyansın tahmininin toplamından elde edilir.

$$v_{Y_i}^* = v_{Y_i} + \tau^2 \quad (2.95)$$

Ağırlıklı ortalama (M^*) ise ağırlıkla etki büyüklüğü çarpımı olan $w_i^* Y_i$ çarpımları toplamının ağırlık toplamına bölünmesi ile hesaplanır.

$$M^* = \frac{\sum_{i=1}^k w_i^* Y_i}{\sum_{i=1}^k w_i^*} \quad (2.96)$$

Meta analizinde birleştirilmiş çalışmalara ait toplam etki büyüklüğünün varyansı, ağırlıkların toplamının tersine, standart hatası da varyansın kareköküne eşittir.

$$v_{M^*} = \frac{1}{\sum_{i=1}^k w_i^*} \quad (2.97)$$

$$SE_{M^*} = \sqrt{v_{M^*}} \quad (2.98)$$

% 95 Güven aralığı ise aşağıdaki gibi formüle edilir (21).

$$M^* \pm 1.96 SE_{M^*} \quad (2.99)$$

Buna göre, rasgele etki modelinde, etki büyüklükleri ve etki büyüklüklerine ait % 95 güven aralıkları aşağıdaki formülasyonlarla verilmektedir. OR ve % 95 güven aralığı;

$$OR = \exp\left[\frac{\sum_{i=1}^k w_i^* \ln(OR)}{\sum_{i=1}^k w_i^*}\right] \quad (2.100)$$

$$\% 95 \text{ Güven Aralığı} = \exp[OR \pm 1.96\sqrt{1/w_i^*}] \quad (2.101)$$

şeklinde hesaplanır. RR ve % 95 güven aralığı;

$$RR = \exp\left[\frac{\sum_{i=1}^k w_i^* \ln(RR)}{\sum_{i=1}^k w_i^*}\right] \quad (2.102)$$

$$\% 95 \text{ Güven Aralığı} = \exp[RR \pm 1.96\sqrt{1/w_i^*}] \quad (2.103)$$

şeklinde hesaplanır. ARR ve % 95 güven aralığı;

$$ARR = \frac{\sum_{i=1}^k w_i^* ARR}{\sum_{i=1}^k w_i^*} \quad (2.104)$$

$$\% 95 \text{ Güven Aralığı} = ARR \pm 1.96 \sqrt{1/w_i^*} \quad (2.105)$$

şeklinde hesaplanır. NNT ve % 95 güven aralığı ise;

$$NNT = \frac{\sum_{i=1}^k w_i^* NNT}{\sum_{i=1}^k w_i^*} \quad (2.106)$$

$$\% 95 \text{ Güven Aralığı} = NNT \pm 1.96 \sqrt{1/w_i^*} \quad (2.107)$$

şeklinde hesaplanır (21).

2.1.3.3. Sabit Etki Modeli ile Rasgele Etki Modelinin Karşılaştırılması

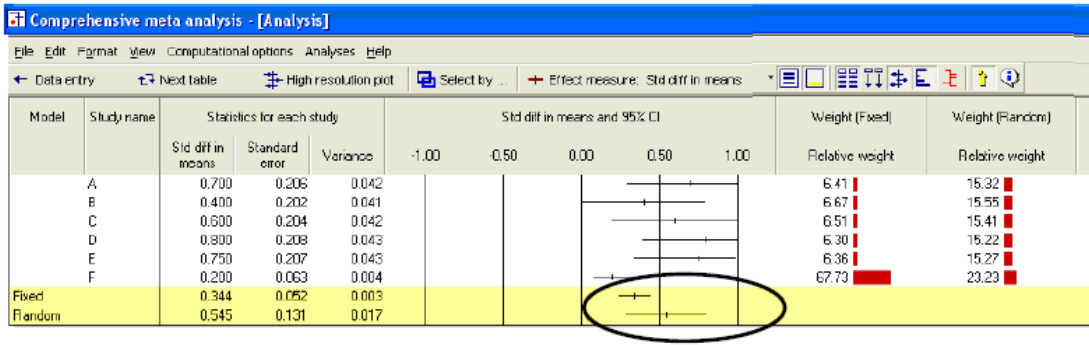
Sabit etki modelinde, gerçek etki büyüklüğünün meta analizine dahil edilen bütün çalışmalarda aynı olduğu varsayılır ve özet etki (summary effect) istatistikleri bu ortak etki büyüklüğünün bir tahminidir. Etki büyüklüklerinin çalışmalar arasında farklılık göstermesinin tek nedeni örnekleme hatasıdır. Bu nedenle farklı çalışmalar ağırlıklandırılırken, büyük örnekleme çalışmalarda aynı etki büyüklüğü hakkında daha fazla bilgi sahibi olduğumuzdan örneklem genişliği küçük olan çalışmalarda çok fazla bilgi ihmal edilmektedir. Buna karşılık, rasgele etki modelinde gerçek etki büyüklüğünün meta analizine alınacak her bir çalışmada aynı olmadığı varsayılır ve bu modelde amaç, etki büyüklüğünü tahmin etmek değil, etki büyüklüklerinin dağılımlarının ortalamasını tahmin etmektir. Özet etki istatistikleri bu etkilerin ortalamasının bir tahminidir. Her bir çalışma farklı etki büyüklüğü hakkında bilgi

sağladığından, tüm bu etki büyüklüklerinin özet tahminlerde gösterildiğinden emin olmak istenir (15).

Bunu bir örnekle açıklamak gerekirse; meta analizine 6 farklı çalışmanın dahil edildiği ve bu çalışmaların (A-E) 5 tanesine ait örneklem büyüklüğünün 100, diğer bir çalışmada (F) örneklem büyüklüğünün 1000 olduğu varsayalım. A-E arasındaki beş çalışmaya ait güven aralıkları geniş ve duyarlılığı (ağırlığı) düşük iken, F çalışması için güven aralığı daha dar ve aynı zamanda duyarlılığı da daha yüksektir (Şekil 2.6). Küçük örneklemlilerle çalışmaların (A-E) etki büyüklükleri 0.40 ile 0.80 arasında değer alırken, büyük örneklemlilerle çalışmanın etki büyüklüğü ise 0.20 gibi bir değere sahiptir. Yani, büyük örneklemlilerle çalışmalar küçük etki büyüklüğüne sahipken, küçük örneklemlilerle çalışmalar daha büyük etki büyüklüklerine sahiptir (15).

Meta analizinde sabit etki modeli tercih edildiğinde, ağırlığın yaklaşık % 68'i F çalışmasına, diğer çalışmaların her birine de yaklaşık olarak % 6'sı atfedilir. Etki büyüklükleri cinsinden incelendiğinde, F çalışması diğer çalışmalara nazaran daha küçük bir etki büyüklüğüne sahiptir ve dolayısıyla sabit etki modelinde meta analizi sonucunda elde edilen tahminler sol tarafta yer almaktadır (15).

Rasgele etki modeli tercih edildiğinde, ağırlığın yaklaşık % 23'ü büyük örnekleme sahip F çalışmasına atfedilirken, diğer küçük örneklemlilerle çalışmaların her birine ise ağırlığın yaklaşık olarak % 15'i atfedilmektedir. Toplam etki büyüklüğü rasgele etki modelinde 0.55 olarak hesaplanırken, sabit etki modelinde ise 0.34 olarak hesaplanmıştır. Büyük örnekleme sahip çalışmanın etkisi rasgele etki modelinde daha az belirgindir. Buradan şunu söyleyebiliriz ki, meta analizine alınacak olan çalışmalarda çok büyük örneklemlilerle çalışmalar varsa ve kullanılan model sabit etki modeli ise toplam etki büyüklüğü büyük örneklemlilerle çalışmanın yönünde olur ve sınırları onlara yakındır. Model olarak rasgele etki modeli tercih edilirse etki büyüklüğünün yönü büyük örneklemlilerle çalışmadan çok küçük örneklemlilerle çalışmaların yönünde olacaktır (15).



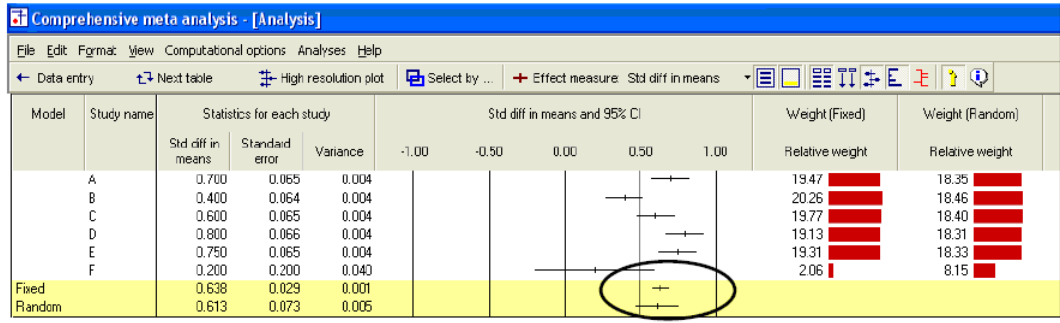
Şekil 2.6: Küçük örneklem büyüklüklerine sahip çalışmaların istatistiksel modellerdeki değişimi

A-E arasındaki 6 çalışmanın büyük örneklemliler olduğu ($n=1000$), F çalışmasının da küçük örneklemliler ($n=100$) bir çalışma olduğu durum incelendiğinde, A-E çalışmaları için güven aralığı dar ve yüksek bir duyarlılığa sahip, F çalışması için güven aralığı oldukça geniş ve daha az duyarlılığa sahip olmaktadır (Şekil 2.7). Büyük örneklemliler çalışmalarının hepsinde etki büyüklüğü 0.40 ile 0.80 arasında değer alırken küçük örneklemliler çalışma da etki büyüklüğü 0.20'dir (15).

Sabit etki modeli tercih edildiğinde, büyük örneklemliler çalışmalarının her biri yaklaşık olarak ağırlığın % 20'si ile küçük örneklemliler çalışma olan F çalışması ise ağırlığın yalnızca % 2'si ile ağırlıklandırılmaktadır. Rasgele etki modeli tercih edildiğinde, büyük örneklemliler çalışmalarının her biri yaklaşık olarak ağırlığın % 18'i ile küçük örneklemliler çalışma olan F çalışması ise ağırlığın yalnızca % 8'i ile ağırlıklandırılmaktadır (15).

Küçük örneklemliler çalışmalarının sabit etki modelinde etkisi hemen hemen yoktur (%2), rasgele etki modelinde etkisi ağırlığın % 8'i kadardır ki bu değer büyük örneklemliler çalışmalarının herhangi birine ait atfedilen ağırlık değerinin yaklaşık yarısı kadardır (15).

Sonuç olarak, rasgele etki modeli ile atfedilen ağırlıklar sabit etki modeli ile atfedilen ağırlıklardan çok daha dengelidir. Çünkü, sabit etki modelinden rasgele etki modeline geçerken ekstrem çalışmalar eğer büyük örneklemliler çalışma ise daha az etkili olacaktır ve eğer küçük örneklemliler çalışma ise daha fazla etkili olacaktır (15).



Şekil 2.7: Büyük örneklem büyüklüklerine sahip çalışmaların istatistiksel modellerdeki değişimi

Varyans, sabit etki modelinde, çalışma içi varyans olarak değerlendirilirken, rasgele etki modelinde çalışma içi ve çalışmalar arası varyansın toplamı olarak değerlendirilmektedir. Aralarındaki bu farklılık, güven aralıklarının genişliklerini etkilemektedir ki böyle bir durumda rasgele etki modeli sabit etki modeline göre daha geniş bir güven aralığına sahiptir (15, 21).

Sabit etki modeli, meta analizine dahil edilen bütün çalışmaların gerçek etki büyüklüklerinin aynı olduğu durumlarda tercih edilmektedir. Ayrıca, çalışmanın amacı, tanımlanmış bir populasyon için ortak bir etki büyüklüğü hesaplamak ve diğer populasyonlar için bir genelleme yapmamaksa böyle bir durumda da sabit etki modeli kullanılmalıdır (21).

Meta analizinde homojenlik testi sonucunda homojen kabul edilen çalışmalar için araştırmacılar sabit etki modelini, heterojen çalışmalar için ise rasgele etki modelini kullanmaktadır. Bazen, homojenlik testi istatistiksel olarak anlamlı bulunmuşsa sabit etki modeli ile başlanıp rasgele etki modeline geçiş uygulaması benimsenmektedir. Bu uygulama kesinlikle tavsiye edilmemelidir, çünkü rasgele etki modelinin kullanılması kararı bütün meta analizine dahil edilen çalışmaların ortak bir etki büyüklüğüne sahip olup olmaması anlayışına dayanmaktadır (21, 38, 39).

2.1.4. Meta Analizinde Heterojenlik

Meta analizine dahil edilen çalışmalar arasındaki değişkenlik heterojenlikle ifade edilmektedir. Heterojenlik genelde, klinik, metodolojik ve istatistiksel olmak üzere üç kategoriye ayrılır. *Klinik heterojenlik (klinik farklılık)*, deneklere, yapılan müdahaleye ve çalışmadaki sonuç değişkeninin çeşitliliğine göre çalışmalardaki klinik farklılık

olarak tanımlanmaktadır. Örneğin, araştırmanın yapıldığı yer, çalışmaya dahil edilen hastaların yaşı, cinsiyeti ve hastalık şiddeti, uygulanacak olan tedavi yöntemi, kullanılacak olan ilaç ve dozu, çalışmadaki sonuç değişkenlerinin farklılıkları klinik heterojenliğin bir göstergesidir. *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 (40,41).

Homojenlik analizi, etki büyüklüklerinin bir çalışmadan diğer bir çalışmaya nasıl değiştiğini gösteren bir ölçüttür. Araştırmacı, homojenlik testi yaparak etki büyüklüğünün varyansının ve örneklemin beklenen hatasının farklılık gösterip göstermediğini incelemeyi amaçlar. Meta analizinde farklı çalışmalardan elde edilen etki büyüklüğü nokta tahminlerinin farklı olması beklenen bir şeydir. Burada dikkat edilecek nokta farklılıkların olup olmaması değil, farklılığın makul bir şekilde görmezden gelinip gelinmeyeceğidir. Bir meta analizinde sonuçların birleştirilmesinden sonra heterojenlik için istatistiksel testlerin ve grafiksel gösterimlerin incelenmesi gerekmektedir. Eğer homojenlik sağlanamamışsa çalışmalar ve sonuçlar arası tutarsızlık varsa böyle bir durumda heterojenlik uygun istatistiksel yöntemlerle analiz yapılarak çalışmanın homojenliği sağlanmalıdır. Aksi takdirde bilgiler istatistiksel açıdan güvenilir olmayabilir (6,8).

Çalışma sonuçlarının güven aralıkları ile verilmesi yani grafiksel olarak ifade edilmesi birleştirilmiş olan etki büyüklüklerinin homojen olup olmadığına karar vermemizi kısmen kolaylaştıracaktır. Güven aralıkları her bir tahminin ne kadar kesin olduğunu ve sonuçların istatistiksel olarak anlamlı olup olmadığını göstermektedir. Eğer tüm çalışmalar gerçekte benzer bir değeri tahmin ediyorsa birleştirilmiş sonuçların dağılım aralığı daha dar olacaktır. Küçük örnekler için güven aralıkları daha geniş olmakla birlikte örnek genişliği artırıldıkça güven aralıkları da daralacaktır (42).

Meta analizi çalışmalarında heterojenlik söz konusu ise böyle bir durumda değişkenliğin iki kaynağı vardır. Bunlardan biri örneklem hatası (çalışma içi değişkenlik, within-study variability), diğeri ise çalışmalar arasındaki değişkenliktir (between-study variability). Meta analizine dahil edilen her bir çalışma farklı örneklemelerden oluştuğu için örneklem hatası yani çalışma içi değişkenlik meta analizinde daima mevcuttur. Çalışmalar arasındaki değişkenlik ise, tek bir çalışma ile

tahmin edilen populasyon etki büyüklükleri arasında gerçekten heterojenlik olduğunda çok açık bir şekilde görülmektedir (31).

Meta analizinde heterojenlik, çalışmalar arası varyans anlamlı derecede arttığında söz konusu olur ki varyansın büyümesinin iki nedeni olabilir. Bunlardan birincisi, çalışmalar arasında kullanılan yöntem farklılığı, diğeri ise çalışmalarda bireysel olarak yapılan analizlerde etki büyüklüğünün farklı alınmış olmasıdır. Heterojenlik testi ve heterojenlik ölçümleri çalışmalar arasındaki varyans değeri ile değil heterojenlikten dolayı artan varyans değeri ile doğrudan ilişkilidir (14,43).

2.1.4.1. Heterojenliğin Testi

2.1.4.1.1. Cochran'ın Q istatistiği

Çalışmalar arasındaki gerçek heterojenliğin olup olmadığını değerlendirmede kullanılan en basit ve çok yaygın olan yaklaşım Cochran (1954) tarafından önerilen (k-1) serbestlik dereceli ki-kare dağılımına sahip Ki-Kare heterojenlik testidir ve bu test Q istatistiği olarak bilinmektedir. Literatürde standart χ^2 testi olarak da geçmektedir (44). $H_0 : \theta_1 = \theta_2 = \dots = \theta_k = \theta$ (tüm çalışmaların etki büyüklükleri eşittir) ya da $H_0 : \sigma_B^2 = 0$ (çalışmalar arası varyanslar eşittir) şeklinde kurulan sıfır hipotezi, “en az bir etki büyüklüğü farklıdır” veya “en az bir varyans farklılık göstermektedir” şeklinde kurulan alternatif hipotezine karşı test edilir. Burada bahsedilen etki büyüklüğü, kullanılan sonuç değişkeninin özet istatistiğine göre değişmektedir. Cochran Q test istatistiği aşağıdaki gibi ifade edilir.

$$Q = \sum_{i=1}^k w_i (Y_i - M)^2 \quad (2.108)$$

M, gözlenen etki büyüklüklerinin ağırlıklı ortalaması olup, Denklem 2.108'de yerine konulduğunda;

$$Q = \sum_{i=1}^k w_i \left(Y_i - \frac{\sum_{i=1}^k w_i Y_i}{\sum_{i=1}^k w_i} \right)^2 \quad (2.109)$$

$$Q = \sum_{i=1}^k w_i \left(Y_i^2 - 2 Y_i \frac{\sum_{i=1}^k w_i Y_i}{\sum_{i=1}^k w_i} + \left(\frac{\sum_{i=1}^k w_i Y_i}{\sum_{i=1}^k w_i} \right)^2 \right) \quad (2.110)$$

$$Q = \sum_{i=1}^k w_i Y_i^2 - 2 \sum_{i=1}^k w_i Y_i \frac{\sum_{i=1}^k w_i Y_i}{\sum_{i=1}^k w_i} + \sum_{i=1}^k w_i \frac{\left(\sum_{i=1}^k w_i Y_i \right)^2}{\left(\sum_{i=1}^k w_i \right)^2} \quad (2.111)$$

$$Q = \sum_{i=1}^k w_i Y_i^2 - 2 \frac{\left(\sum_{i=1}^k w_i Y_i \right)^2}{\sum_{i=1}^k w_i} + \frac{\left(\sum_{i=1}^k w_i Y_i \right)^2}{\sum_{i=1}^k w_i} \quad (2.112)$$

$$Q = \sum_{i=1}^k w_i Y_i^2 - \frac{\left(\sum_{i=1}^k w_i Y_i \right)^2}{\sum_{i=1}^k w_i} \quad (2.113)$$

şekline dönüşür ve çalışmaların ortalamaları arasındaki farklılığın karelerinin ağırlıklı toplamıdır. Elde edilen değer, (k-1) serbestlik dereceli ki-kare dağılımına sahip istatistik değeri ile karşılaştırılarak test edilir (21, 40, 45, 46).

Q test istatistiğinin beklenen değeri;

$$E[Q] = \sigma_B^2 \left(\sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i} \right) + (k-1) \quad (2.114)$$

şeklinde hesaplanır. Q eşitliği ile çalışmalar arası varyans (σ_B^2) için ya bir aralık tahmini (moment estimate) kullanılarak ya da maksimum likelihood tahmini (maksimum olabilirlik tahmini) yöntemi kullanılarak aşağıdaki formülasyonla elde edilir (43, 45).

$$\hat{\sigma}_B^2 = \frac{Q - (k-1)}{\sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i}} \quad (2.115)$$

Eğer $Q < (k-1)$ ise çalışmalar arası varyansın tahmini değeri olan $\hat{\sigma}_B^2$ negatif bir değer alır. Böyle bir durumla karşılaşılığında çalışmalar arası varyansın tahmini değeri $\max(0, \hat{\sigma}_B^2)$ gibi kullanılır. Böylece, rasgele etki modeline ait etki büyüklüğünün duyarlılığı (ağırlığı) $(1/v_R)$, sabit etki modeline ait etki büyüklüğünün duyarlılığından $(1/v_F)$ daha büyük olmayacaktır (43).

Çalışma içi varyansın tahmini ise, iki farklı olasılıkla elde edilir. Birincisi, Takkouche ve ark. (1999)'nın belirttiğine göre aşağıdaki şekilde hesaplanmaktadır (47).

$$\hat{\sigma}_{w,1}^2 = kv_F = \frac{k}{\left(\sum_{i=1}^k w_i\right)} \quad (2.116)$$

İkinci olasılık ise,

$$\hat{\sigma}_{w,2}^2 = \frac{(k-1)\sum_{k=1}^k w_i}{\left(\sum_{i=1}^k w_i\right)^2 - \sum_{i=1}^k w_i^2} \quad (2.117)$$

şeklinde formüle edilmektedir. Çalışma içi varyansın tahminine dayalı Q test istatistiğinin beklenen değeri ise aşağıdaki gibi hesaplanır (43).

$$E[Q] = (k-1) \left(\left(\frac{\sigma_B^2}{\sigma_{w,2}^2} \right) + 1 \right) \quad (2.118)$$

Çalışma içi varyans tahmini $\sigma_{w,2}^2$, hem $E[Q]$ hem de heterojenlik testinin gücü ile doğrudan ilişkilidir. Yani, $\sigma_{w,2}^2$ 'nin kullanımı ya heterojenliği test etmede ya da heterojenlik ölçüsü miktarını belirlemede heterojenlik etkisini daha iyi ifade etmektedir.

Bu nedenle, $\hat{\sigma}_{w,2}^2$, heterojenlik testinin gücü üzerine çalışma içi varyansın etkisini değerlendirmek konusunda $\hat{\sigma}_{w,1}^2$ 'e tercih edilir. Eğer $w_1 = w_2 = \dots = w_k$ ise $\hat{\sigma}_{w,1}^2 = \sigma_{w,2}^2$ olduğuna dikkat etmek gerekir (43).

Heterojenliğin test edilmesi, Cochran Q test istatistiğinin düşük güce sahip olmasından ya da heterojenlik ölçümü miktarının önemsiz bulunmasından dolayı yanlış anlaşılmaktadır. Q istatistiğinin gücü meta analizine dahil edilen çalışma sayısından etkilenmekte ve çalışma sayısı arttıkça Q değeri de artmaktadır. Meta analizine dahil edilen çalışma sayısı az ise testin gücü düşük, fazla ise testin gücü yüksek olmaktadır. Baujat ve ark. (2002) anlamlılık testi sonucunun heterojenlik için kuvvetli bir kanıt olarak değerlendirilmesi gerektiği üzerinde durmaktadır (48). Ancak bu durum meta analizine dahil edilecek olan çalışmaların sayısı az olduğunda doğrudur. Eğer meta analizine pek çok çalışma dahil edilmişse, meta analizi için önemsiz olan küçük bir miktardaki heterojenlik istatistiksel olarak anlamlı bulunabilir. Yani pek çok çalışma, meta analizine dahil edildiğinde, heterojenlik miktarı çok küçük de olsa bu değer test edildiğinde istatistiksel olarak anlamlı kabul edilebilir. Böyle bir durumda heterojenliğin olup olmamasına da klinik olarak önemli olup olmamasına da karar vermek zorlaşacaktır. Bu yüzden sadece istatistiksel test ile değerlendirmek değil aynı zamanda meta analizi üzerinden heterojenliğin etkisini de tanımlanması tavsiye edilmektedir. Yani, test istatistiği olan Q, heterojenlik ölçüsü gibi tek başına kullanılmamalıdır, heterojenlik ölçüsü değerleri de hesaplanmalıdır. (31, 43, 46, 49).

2.1.4.2. Heterojenlik Ölçümleri

Heterojenliğin varlığında heterojenliğin miktarının belirlenmesi için literatürde pek çok heterojenlik ölçümü verilmektedir. H^2 , R^2 , τ^2 ve I^2 istatistikleridir ki bunlardan ilk ikisinin karekökü alınarak elde edilmektedir (43, 46).

2.1.4.2.1. H istatistiği

Higgins ve Thompson tarafından (2002) elde edilen bir indekstir ve bu indeks Birge'nin oranı olarak da adlandırılmaktadır. Cochran'ın Q istatistiğinden yararlanarak aşağıdaki gibi tanımlanmaktadır (50).

$$H^2 = \left\{ \begin{array}{ll} \frac{Q}{(k-1)}, & Q > df \\ 1 & , \quad Q \leq df \end{array} \right\} \quad (2.119)$$

H^2 değeri 1 ile ∞ arasında bir değer almaktadır. H 'nin 1 olması homojenliğin mükemmel olduğunu göstermektedir ve çalışma sayısı ile sistematik olarak bir artış göstermektedir. Bu değer beklenen değeri ise aşağıda gibi verilmektedir (31, 43, 46).

$$E[H^2] = \frac{E[Q]}{(k-1)} = \frac{\sigma_B^2}{\sigma_{W,2}^2} + 1 = \frac{\sigma_B^2 + \sigma_{W,2}^2}{\sigma_{W,2}^2} \quad (2.120)$$

H istatistiğine ait % 95 güven aralığı aşağıdaki şekilde hesaplanmaktadır:

$$\exp[\ln(H) \pm 1.96 * SE(\ln(H))] \quad (2.121)$$

$SE(\ln(H))$ için, Cochran Q istatistiğinin çalışma sayısından büyük veya küçük oluşuna göre farklı formüllerle hesaplanmaktadır. Buna göre, $Q > k$ olduğu durumda,

$$SE(\ln(H)) = \frac{1}{2} \left(\frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{(2k-3)}} \right) \quad (2.122)$$

şeklinde, $Q \leq k$ olduğu durumda,

$$SE(\ln(H)) = \sqrt{\frac{1}{2(k-2)} \left(1 - \frac{1}{3(k-2)^2} \right)} \quad (2.123)$$

şeklinde hesaplanmaktadır (31,50). H 'nin değeri çalışma sayısı küçük olduğunda ($k < 8$) küçük bir yanlılık gösterebilir. H 'nin değişkenliği çalışma sayısı az olduğunda büyüktür, böylece pratikte orta derecede olan bir heterojenliği bile ayırt etmek zorlaşabilmektedir (% 50 şansla). Çalışma sayısı artırıldığında değişkenlik yavaşça

azalacaktır. Ayrıca, homojen durumda H'nin güven aralığı mutlaka 1 değerini içermelidir. Aksi takdirde heterojenliğin varlığından bahsedilmektedir (50).

H istatistiği heterojenlikten dolayı varyans şişirme faktörü olarak yorumlanmaktadır. Pratikte $\max(1, H^2)$ kullanılır. H^2 'nin düzenlenmiş hali ise aşağıdaki gibi verilmektedir (43).

$$H_M^2 = \begin{cases} \frac{Q - (k - 1)}{(k - 1)}, & Q > df \\ 0, & Q \leq df \end{cases} \quad (2.124)$$

Homojenite durumunda sıfır değerini almakta ve pratikte $\max(0, H_M^2)$ kullanılmaktadır. Buradan da anlaşıldığı gibi 0 ile ∞ arasında bir değer almaktadır. Beklenen değeri ise, çalışmalar arası varyansın çalışma içi varyansa oranı ($E[H_M^2] = \frac{\sigma_B^2}{\sigma_{W,2}^2}$) olarak hesaplanmakta ve bu yolla ortalama H_M^2 üzerinden çalışma içi varyans çalışmalar arası varyansla ilişkisi verilmektedir (43).

2.1.4.2.2. τ^2 istatistiği

Çalışmalar arası varyansdır ve meta analizindeki rasgele etki modelindeki gibi tahmin edilir. Meta analizinde τ^2 'nin tahmini için pek çok öneri vardır. τ^2 'yi tahmin etmek için DerSimonian ve Laird yönteminden yararlanılır ve aşağıdaki şekilde hesaplanır (21):

$$\tau^2 = \frac{Q - df}{C} \quad (2.125)$$

Formülde yer alan C aşağıdaki gibi hesaplanmaktadır.

$$C = \sum_{i=1}^k w_i - \frac{\sum_{i=1}^k w_i^2}{\sum_{i=1}^k w_i} \quad (2.126)$$

C formülasyonda yerine konursa çalışmalar arası varyans aşağıdaki gibi tanımlanır.

$$\hat{\tau}^2 = \begin{cases} \frac{Q - (k-1)}{\sum w_i - \frac{\sum w_i^2}{\sum w_i}} & Q > (k-1) \\ 0 & Q \leq (k-1) \end{cases} \quad (2.127)$$

τ^2 'nin dağılımı normal dağılıma uygun değildir. Bu değer için % 95 güven aralığını elde edebilmek için pek çok yöntem vardır ve bunların içerisinde en basit olanı aşağıda verilmektedir. Alt sınır değeri,

$$L = \exp(0.5 * \ln(\frac{Q}{k}) - 1.96 * B) \quad (2.128)$$

şeklinde, üst sınır değeri,

$$U = \exp(0.5 * \ln(\frac{Q}{k}) + 1.96 * B) \quad (2.129)$$

şeklinde hesaplanır. Denklem 2.128 ve Denklem 2.129'da yer alan B, Cochran Q istatistiğinin çalışma sayısından büyük veya küçük oluşuna göre farklı formüllerle hesaplanmaktadır. Eğer $Q > k$ ise,

$$B = 0.5 * \frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{(2k-3)}} \quad (2.130)$$

şeklinde, $Q \leq k$ ise,

$$B = \sqrt{\frac{1}{2(k-2) \left[1 - \left(\frac{1}{3(k-2)^2} \right) \right]}} \quad (2.131)$$

şeklinde hesaplanmaktadır. % 95 Güven aralığına ait alt sınır değeri,

$$LL_{\tau^2} = \frac{(k-1)*(L^2-1)}{C} \quad (2.132)$$

şeklinde, üst sınır değeri ise,

$$UL_{\tau^2} = \frac{(k-1)*(U^2-1)}{C} \quad (2.133)$$

şeklinde hesaplanmaktadır (21).

τ^2 , alt sınır veya üst sınır değerlerinden herhangi biri “0”dan küçük ise gerçek değeri yerine 0 değeri atanır. Eğer alt sınır değeri 0’den büyük ise τ^2 istatistik olarak anlamlı kabul edilir. Bununla birlikte, τ^2 değeri Q temelli olduğundan ve Q’nun örnekleme dağılımı daha iyi bilindiğinden anlamlılık için tercih edilecek yöntem Q testi olmalıdır ve H_0 hipotezi reddedildiğinde yani heterojen olduğunda kullanılacak olan heterojenlik ölçümü τ^2 olmalıdır (21).

2.1.4.2.3. R istatistiği

Sabit etki modeline nazaran rasgele etki modeline ait güven aralıklarının etkisini tanımlamakta kullanılır. Meta analizine dahil edilecek olan çalışma sayısından bağımsızdır. H istatistiğine benzerdir ve τ^2 ’dan hesaplanır ve aşağıdaki şekilde formüle edilir (46, 50).

$$R^2 = \frac{\tau^2 + \sigma^2}{\sigma^2} \quad (2.134)$$

R=1 olması homojenliğin mükemmel olduğunu göstermektedir. R’nin değeri ve % 95 güven aralıkları hesaplamaları ve yorumlamaları H istatistiği ile çok benzerdir. Buna göre R’ye ait % 95 güven aralığı aşağıdaki şekilde hesaplanmaktadır.

$$\exp[\ln(R) \pm 1.96 * SE(\ln(R))] \quad (2.135)$$

SE(ln(R)) için, Cochran Q istatistiğinin çalışma sayısından büyük veya küçük oluşuna göre farklı formüllerle hesaplanmaktadır. Buna göre, $Q > k$ olduğu durumda aşağıdaki gibi hesaplanmaktadır.

$$SE(\ln(R)) = \frac{1}{2} \left(\frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{(2k-3)}} \right) \quad (2.136)$$

$Q \leq k$ için;

$$SE(\ln(R)) = \sqrt{\frac{1}{2(k-2)} \left(1 - \frac{1}{3(k-2)^2} \right)} \quad (2.137)$$

şeklinde hesaplanmaktadır (50).

2.1.4.2.4. I^2 İndeksi

Son yıllarda, Higgins ve Thompson (2002) tarafından önerilen ve Cochran'ın Q indeksinden ve H^2 indeksinden yararlanarak oluşturulmuş bir indeks olup aşağıdaki gibi farklı şekillerde formüle edilmektedir (31, 46).

$$I^2 = \begin{cases} \frac{Q - (k-1)}{Q} * 100\%, & Q > (k-1), \\ 0, & Q \leq (k-1) \end{cases} \quad (2.138)$$

$$I^2 = \begin{cases} \frac{H^2 - 1}{H^2} * 100\%, & Q > (k-1), \\ 0, & Q \leq (k-1) \end{cases} \quad (2.139)$$

$$I^2 = \begin{cases} \frac{c\tau^2}{Q} * 100\%, & Q > (k-1), \\ 0, & Q \leq (k-1) \end{cases} \quad (2.140)$$

I^2 'nin beklenen değeri;

$$E(I^2) = \frac{\tau^2}{(\tau^2 + \sigma^2)} \quad (2.141)$$

şeklinde hesaplanmaktadır. $\tau^2=0$ için Q yaklaşık olarak (k-1) serbestlik dereceli ki-kare dağılımı göstermekte ve $1/Q$ 'nun beklenen değeri $df > 2$ olacak veya $k > 3$ olacak şekilde $1/(df-2)$ olasılıkla ters ki-kare dağılımı göstermektedir. Bu yüzden I^2 'nin beklenen değeri $\tau^2=0$ ve $k > 3$ için $E[I^2] = 100 * \left(\frac{-2}{(k-3)} \right)$ şeklinde hesaplanmaktadır.

Bu olasılık yalnızca en az 4 çalışma için tanımlanmakta ve meta analizine dahil edilen çalışma sayısına bağlı olmaktadır, bununla birlikte eğer çalışma sayısı artırılırsa bu bağımlılık daha küçük olacaktır (43). I^2 için % 95 güven aralığına ait alt sınır değeri,

$$L = \exp(0.5 * \ln(\frac{Q}{k-1}) - 1.96 * B) \quad (2.142)$$

şeklinde, üst sınır değeri,

$$U = \exp(0.5 * \ln(\frac{Q}{k-1}) + 1.96 * B) \quad (2.143)$$

şeklinde hesaplanmaktadır. Denklem 2.142 ve Denklem 2.143'de yer alan B, Cochran Q istatistiğinin çalışma sayısından büyük veya küçük oluşuna göre farklı formüllerle hesaplanmaktadır. Eğer $Q > k$ ise,

$$B = 0.5 * \frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{(2k-3)}} \quad (2.144)$$

Eğer $Q \leq k$ ise,

$$B = \sqrt{\frac{1}{2(k-2) \left[1 - \left(\frac{1}{3(k-2)^2} \right) \right]}} \quad (2.145)$$

şeklinde hesaplanır. % 95 Güven aralığına ait alt sınır değeri,

$$LL_{I^2} = \frac{(L^2 - 1)}{L^2} * 100\% \quad (2.146)$$

şeklinde, üst sınır değeri,

$$UL_{I^2} = \frac{(U^2 - 1)}{U^2} * 100\% \quad (2.147)$$

şeklinde hesaplanmaktadır. I^2 , alt sınır veya üst sınır değerlerinden herhangi biri “0”dan küçük ise bu değere 0 değeri atanır. Eğer alt sınır 0’dan büyük ise I^2 istatistik olarak anlamlıdır. Bununla birlikte, I^2 değeri Q temelli olduğundan ve Q’nun örnekleme dağılımı daha iyi bilindiğinden anlamlılık için tercih edilecek yöntem Q testi olmalıdır ve H_0 hipotezi reddedildiğinde yani heterojen olduğunda kullanılacak olan heterojenlik ölçümü I^2 olmalıdır (21).

Heterojenlik ölçüsü bir ölçek üzerinde doğal olarak 0 ile % 100 arasında görülmektedir. % 100’e yakın değerler için heterojenlik çok büyük, sıfıra yakın değerler için ise küçük heterojenlik şeklinde değerlendirilir. I^2 değeri % 100’e asla ulaşmaz ve çok nadiren % 90’nın üzerinde bir değer alır (37). Fakat bu iki uç değer arasındaki değerler nasıl yorumlanmalıdır? Higgins ve Thompson (2002), I^2 değeri için üç kategori belirlemiştir ve % 25 düşük, % 50 orta ve % 75 ise yüksek heterojenliği ifade etmektedir. % 50’nin üzerindeki değerler için heterojenliğin önemli olduğu fark edilmektedir (43). Patsopoulos ve ark. (2008), % 25’in altında ise düşük, %25-50 arası orta heterojenlik, %50’nin üzerinde olduğunda ise heterojenliğin yüksek olduğu şeklinde yorumlamaktadır (51).

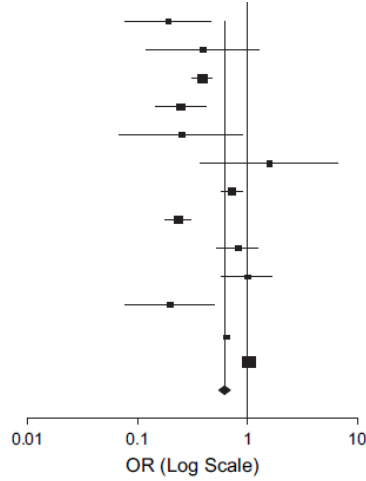
2.1.4.3. Grafik gösterimleri

Bir meta analizinde, birleştirilmiş olan bulguların homojen olup olmadığı, Cochran Q heterojenlik testi, heterojenlik ölçümleri ve grafik gösterimleri ile test edilmektedir. Grafik gösterimlerinde, çalışma sonuçlarının güven aralıkları ile verilmesi her bir tahminin ne kadar kesin olduğunu ve sonuçların istatistiksel olarak anlamlı olup olmadığını daha net bir şekilde ortaya koymaktadır. Heterojenliği ortaya koymada çok etkili olan grafik türleri aşağıdaki gibi sıralanmaktadır (52).

1. Forest grafik
2. Galbraith Radial grafik
3. L'abbe grafik
4. Funnel grafik
5. Normal quantile grafik
6. Standardized residual histogram

2.1.4.3.1. Forest grafik

Meta analizinde çok yaygın olarak kullanılan bir grafik türüdür. Birçok kaynakta meta analiz diyagramı olarak adlandırılmaktadır. Sonuç değişkenimiz binary yapıda olup OR ya da RR ise genellikle etki büyüklüklerinin logaritmik ölçüm değerleri kullanılmaktadır. Grafikte y eksenini meta analizine dahil edilen çalışmaları x eksenini ise OR veya RR'ye ait logaritmik değeri göstermektedir (Şekil 2.8). Bu grafik, bireysel çalışmalar için ve özet istatistikleri için nokta tahminlerini ve % 95 güven aralıklarını göstermek amacıyla kullanılmaktadır. Bu grafik ile her bir bireysel çalışmadan elde edilen tahminler arasındaki değişkenlik, denemeler arasındaki heterojenliği vurgulamaktadır (40, 53, 54).



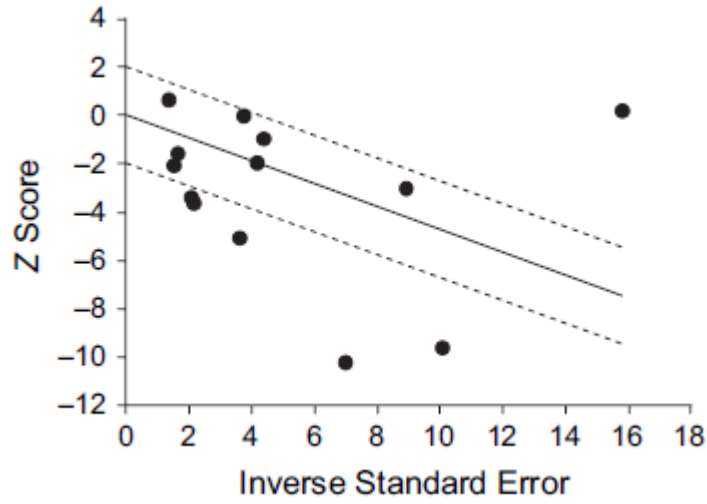
Şekil 2.8: Forest grafik

2.1.4.3.2. Galbraith Radial grafik

Galbraith grafik, Galbraith radial grafik ya da radial grafik olarak da adlandırılmaktadır. Galbraith Radial grafik, ilk kez Galbraith (1988) tarafından geliştirilmiş olup heterojenliği ortaya koymak için kullanılan bir grafikdir. Grafik, her bir çalışmadan elde edilen sonuçların varyansın kareköküne bölünmesinden elde edilen değerler ($\frac{Ln(OR)}{SE}$) (z istatistiği olarak bilinir) standart hatanın tersine karşılık çizilir.

Yatay eksen de $\frac{1}{SE}$ ve düşey eksen de her bir z istatistiği yer almaktadır. Eğer yatay eksenindeki değişken sıfır değerini alıyorsa standart hata sonsuz değerini alacaktır ve böylece düşey eksen de (paydada standart hata olduğundan dolayı) sıfır değerini alacaktır.

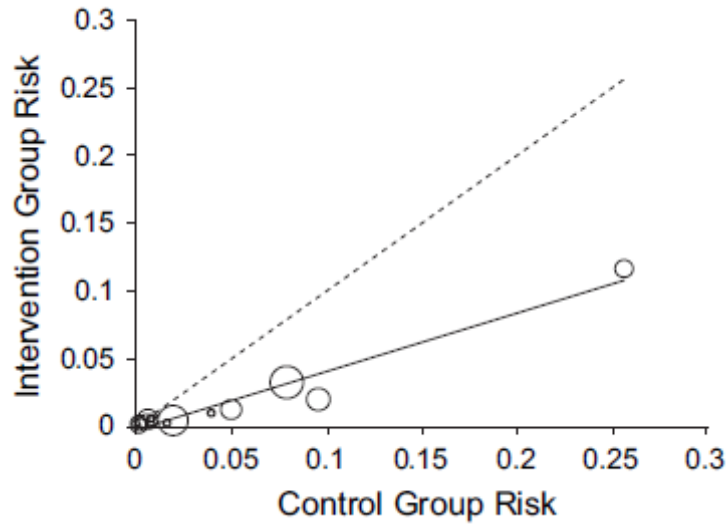
Meta analizinde standart hatası büyük, ağırlığı küçük olan noktalar y eksenine yaklaşır. Orjinden geçen doğrudan gözlenen eğim, çalışmada tahmin edilen sonuca karşılık gelir. Şekil 2.9'da gözlenen eğim ve % 95 güven aralıkları yer almaktadır. Noktalar %95 güven aralığının sınırları içerisinde yer alıyorsa meta analizine dahil edilen çalışmaların homojen olduğu yani benzer çalışmalar olduğu sonucuna varılır. Şekil 2.9'da % 95 güven aralıklarının dışında 6 nokta yer almaktadır. Buna göre, meta analizine dahil edilen çalışmaların heterojen olduğu söylenebilmektedir (52, 54).



Şekil 2.9: Galbraith radial grafik

2.1.4.3.3. L'abbe grafik

Labbe ve ark. tarafından 1978 yılında geliştirilmiş bir grafikdir ve yalnızca sonuç değişkeni binary yapıda olan çalışmaların meta analizlerinde uygulanır. Kontrol ve tedavi grubu gibi iki gruba ait risk oranlarından yararlanılarak çizilen bir grafik türüdür. Şekil 2.10'da yatay (x) ekseninde kontrol grubuna ait risk değerleri, düşey (y) ekseninde ise tedavi grubuna ait risk değerleri yer almaktadır ve her bir gruptaki hesaplanan risk değerlerini (OR) gösteren ortada köşegen çizgi ve regresyon çizgisi yer almaktadır. Noktaların boyutları çalışmanın ağırlığı ile orantılıdır (40, 54).



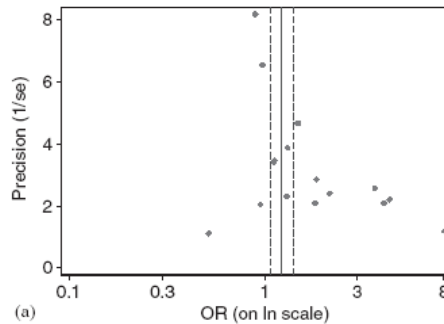
Şekil 2.10: L'abbe grafik

Eğer meta analizine dahil edilen çalışmalar homojen ise noktalar çizgiye yakın, heterojen ise çizgiden büyük sapmalar veya saçılımlar söz konusu olacaktır. Doğrunun altında kalan noktalar, tedavi grubundaki etki büyüklüğü değerinin (riskin) kontrol grubundaki etki büyüklüğü değerinden (riskinden) daha üstün olduğunu ve bu nedenle yeni tedavinin standart tedaviden ya da kontrolden daha az etkili olduğunu göstermektedir. Doğrunun üstünde olan noktalar ise kontrol grubu ile tedavi grubu karşılaştırıldığında yeni tedavinin etkinliğini göstermektedir (52).

2.1.4.3.4. Funnel grafik

Light ve Pillemer tarafından 1985 yılında geliştirilmiş bir grafik türüdür. Funnel grafik (huni grafik), meta analizlerinde tamamlayıcı çalışmaların saçılım grafiğidir. Yatay ekseninde (x ekseninde) etki büyüklüğü ($\ln(OR)$), dikey ekseninde (y ekseninde) ise standart hatanın tersi ($1/SE$), standart hatanın kendisi, örneklem büyüklüğü ya da ağırlık yer almaktadır. Her bir çalışma eşit büyüklükteki noktalarla gösterilir. Funnel grafiğin basitliği, kolaylığı ilgi çekmesine rağmen istatistiksel olarak test edilebilecek belli bir modele dayanmamaktadır. Sonuç olarak, çeşitli istatistiksel yöntemler grafiğin kişisel olarak görsel bir şekilde yorumlanabilmesi için geliştirilmektedir (52).

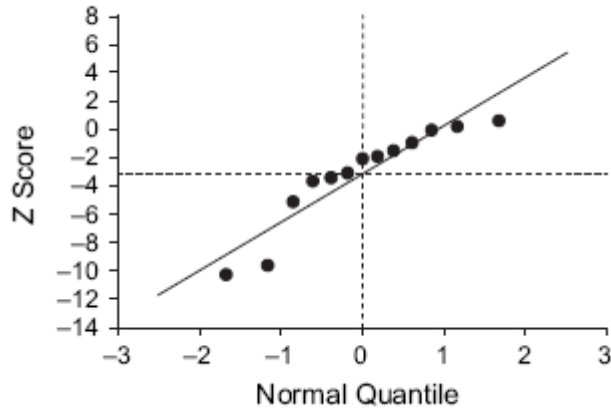
Şekil 2.11’de düz çizgiler etki büyüklüğü değerlerini, kesik çizgiler ise % 95 güven aralıklarını göstermektedir. Bu grafiğin funnel’ın alt sol taraftan missing çalışmaları ile asimetrik olduğu görülmektedir. Asimetrikliğe çalışmalar arasındaki heterojenlik ve çalışmanın kalitesi dahil olmak üzere pek çok faktör neden olabilir ve herhangi böyle bir asimetrikliğin gerçek nedeni ortaya çıkarmak mümkün olamamasına rağmen burada bunun nedeninin yayın yanlılığı (publication bias) olduğu varsayılmaktadır (55).



Şekil 2.11: Funnel plot

2.1.4.3.5. Normal quantile grafik

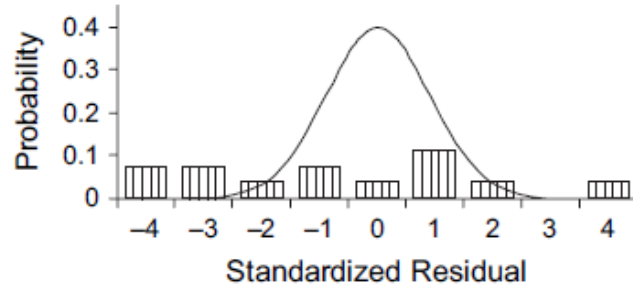
İstatistikte, genellikle verilerin normal dağılıp dağılmadığını kontrol etmekte kullanılmaktadır. Bu grafik meta analizinde Wang ve Bushman (1998) tarafından ilk kez ortaya atılmış olup, y ekseninde her bir çalışmaya ait etki büyüklüklerinin standartlaştırılmış değeri (z puanı), x ekseninde ise her bir çalışmanın rankının standartlaştırılmış yüzdelikleri yer almaktadır (Şekil 2.12). Bu grafik türü, heterojenliği araştırmak ve yayın yanlılığının varlığını değerlendirmek ve verilerin normal dağılıp dağılmadığını kontrol etmek için meta analizi çalışmalarında kullanılmaktadır. Bu grafik türü bazı durumlarda funnel grafiğe alternatif olarak geliştirilmiştir. Özellikle % 95 güven aralıklarını içine aldığından Funnel grafiğe göre yorumlamak daha kolay olmaktadır. Çoğu istatistikçiler bu grafiği scatter grafiğin özel bir türü olarak tanımlamaktadırlar. Grafiğe göre normal dağılımdan bahsedebilmek için noktaların doğru çizgisi üzerinde yer alıyor olması, heterojenlikten bahsedebilmek için ise noktaların kümeleniyor olması gerekmektedir (54, 56).



Şekil 2.12: Normal quantile grafik

2.1.4.3.6. Standardize edilmiş artık histogramı (Standardized residual histogram)

Bu grafik ilk olarak Greenland (1987) tarafından, ardından da Sutton ve ark. (2000) tarafından meta analizi çalışmalarında heterojenliği değerlendirmek ve normallikten sapmaları ortaya koymak amacıyla geliştirilmiştir (53).



Şekil 2.13: Standardize edilmiş artık histogramı

2.1.4.4. Meta Analizindeki Heterojenliğin Nedenleri

Meta analizinde heterojenlik söz konusu ise bu heterojenliğin etkisini ortaya koymak veya heterojenliğin olası nedenlerini tanımlamak çok önemlidir. Meta analizine dahil edilen çalışmaların sonuçlarının arasında şansa atfedilebilenden daha fazla heterojenliğin varlığı sonucuna varılırsa, sonraki uygulanacak adım, var olan farklılıkların kaynağını araştırmaktır.

Şans, tedavi etkilerinin tanımlanması, bazı deneme özellikleri ve diğer faktörler meta analizindeki heterojenliğin olası nedenleridir (57). Şans veya rasgele örnekleme hatasından dolayı farklı çalışmaların sonuçları bütün çalışmalar etki büyüklüğü altında aynı hesaplandığında bile farklı olacaktır. Örneğin rasgele değişkenlik çok küçük çalışmalarda oldukça büyür ve çok geniş çalışmalarda ise daha küçük bir problemdir. Eğer bir meta analizi, oldukça fazla sayıda çalışmadan oluşuyorsa ve çalışmalar içindeki farklılık yalnızca rasgele değişkenlikten dolayı ise çalışma sonuçları ortalama etrafında dağılacaktır ve sonuçlar ortalamadan gittikçe uzaklaşırsa daha az sayıda çalışmadan oluşacaktır. İstatistiksel yöntemler rasgele hatanın büyüklüğü tahmin etmekte kullanılabilir.

Meta analizinde çalışmalar içindeki gözlenen heterojenlik, genellikle yalnızca random değişkenlikle ifade edilemez. Çalışmaya dâhil edilen deneklerin özellikleri ve müdahaleleri heterojenliğin klinik olarak anlamlı nedenleridir. Farklı çalışmaların sonuçları farklı olabilir çünkü farklı çalışmalardaki çalışmaya alınan denekler farklıdır ve bu yüzden tedaviye farklı cevap verirler. Heterojenliğin nedenleri olarak deneklerin özelliklerini tanımlamak (hastalığın şiddeti, tanı, yaş, cinsiyet vb) tedaviden daha az

veya daha fazla yararlı olabileceğini teşhis etmek için klinik de önemlidir. Bu hastalara uygulanan tedavinin daha iyi uyarlanmasına izin verebilir (57).

Benzer olarak, zaman, ortam ve müdahalelerdeki çeşitlilik çalışmalar içindeki sonuçlardaki farklılıklar oldukça fazla olabilir. Örneğin, hastaya bakan kişinin eğitimi düzeyi farklı olabilir. Alternatif olarak, hem tedavi hem de kontrol grubundaki kişilerin aldıkları ilacın dozu etkinin tahmininde değişkenliğin bir nedeni olabilir.

Hastanın atanma yöntemleri, körlüme ve izlem süresi (length of follow-up) gibi deneme düzeni faktörleri çalışma sonuçlarındaki farklılığın önemli nedenleri de olabilir. Çalışma deneme faktörleri heterojenliğin metodolojik nedenleri önemlidir ve bazen klinik olarak anlamlı da olabilir. Örneğin, izlem süresinin periyotlarının farklı olduğu çalışmalar farklı sonuçlar verebilir (57).

2.1.4.5. Meta Analizindeki Heterojenliğin Açıklanması

Meta analizinde heterojenliğin varlığında, sonuçların yorumlanması ve değerlendirilmesi aşamalarının nasıl yapılacağı ile ilgili tam bir fikirbirliği bulunmamaktadır.

Leandro'a (2005) göre, meta analizine dahil edilen çalışmalar arasında farklılık anlamlı değilse yani çalışmalar homojen ise daima sabit etki modeli kullanılmalıdır. İstatistiksel heterojenlik anlamlı ise yani heterojenlik varsa, asla görmezden gelinmemelidir. Heterojenliğin kaynakları, nedenleri açıklanmalıdır. Heterojenliğin varlığında, uç (outlier) çalışmaların tespit edilmesi gereklidir ve analizler uç çalışmaların olup olmamasına göre değerlendirilmelidir. Analiz aşamasında, çalışmalar arasında daha fazla homojenlik sağlamak ve heterojenliği ortadan kaldırmak amacıyla uç çalışmaların analiz dışında tutulması önerilmektedir. Uç çalışmaların analiz dışında bırakılması durumunda, bazı analizlerde homojenlik sağlanırken, bazılarında homojenliğin hala sağlanamadığı gözlenebilmektedir. Heterojenliğin nedenleri, çalışmalar çok dikkatli incelenmesine rağmen yine de tespit edilemiyorsa, böyle bir durumda rasgele etki modeli kullanılmalıdır. Deneklerin karakteristik özelliklerinde, dahil etme ve hariç tutma kriterlerindeki olası farklılığı tanımlamak için çalışmalar farklı homojen gruplar içerisinde sınıflandırılabilir. Her bir homojen alt grup içinde öncelikle heterojenlik testi yapılır. Alt gruplarda homojenlik sağlanabilmişse sabit etki

modeli tercih edilmelidir, şayet homojenlik sağlanamamışsa böyle bir durumda da rasgele etki modeli kullanılmalıdır. Kısaca, çalışmalar arasında heterojenliği açıklayabilmek için, iki önemli yöntem olan alt grup analizleri ve meta regresyon yöntemleri kullanılmalıdır (14, 58, 59).

3. GEREÇ VE YÖNTEM

Meta analizinde çalışmaya dahil edilen her bir çalışma sayısal olarak birleştirilmeden önce çalışmaların her biri için sonuç değişkenine ait ölçümlerin (etki büyüklüklerinin) hesaplanması gerekmektedir. k çalışmaya ait meta analizi değerlendirilirken, meta analizine dahil edilen her bir çalışmanın sonuç değişkeni binary yapıda ise yaygın kullanılan etki büyüklükleri RR, OR, ARR ve NNT özet istatistikleridir.

Bu çalışmada, meta analizine dahil edilen her bir çalışmanın sonuç değişkeninin binary yapıda olduğu, binary yapıda kullanılan etki büyüklüğü OR değerinin beklenen değerinin 1'e eşit olduğu durumlar göz önünde bulundurularak simülasyonlar iki bölümde ele alınmıştır. İlk bölümde, meta analizine dahil edilen çalışmaların homojen olduğu, ikinci bölümde de meta analizine dahil edilen çalışmaların heterojen olduğu durumlara ait, farklı örneklem büyüklükleri ve tedavi etkinlik oranlarında veriler üretilmiş ve bu verilere ait Cochran Q heterojenlik testi, testin anlamlılık değeri ve heterojenlik ölçümleri hesaplatılmıştır. Bu simülasyon sonuçlarından yararlanarak, heterojenlik testi ve heterojenlik ölçümlerinin, meta analizine dahil edilen çalışma sayısından ve örneklem büyüklüğünden, her iki gruba ait tedavi etkinlik oranları arasındaki farktan etkilenip etkilenmedikleri ortaya konmaya çalışılmıştır. Ayrıca, heterojenlik testine ait anlamlılık seviyesi için ve heterojenlik ölçümleri için ROC analizleri yapılarak en uygun kesim noktaları elde edilmeye çalışılmıştır.

3.1. Meta Analizine Dahil Edilen Çalışmaların Homojen Olduğu Durumlar

3.1.1. Simülasyon Çalışması

Tedavi ve kontrol grupları için tedavi etkinlik oranının her iki grupta birbirine eşit (homojen) olduğu bazı kombinasyonlar (0.05; 0.10; 0.20; 0.50; 0.80; 0.90 ve 0.95) göz önünde bulundurularak MINITAB 15.0 (59) paket programının demo versiyonunda örneklem büyüklükleri eşit olmak üzere binomiyal dağılım gösteren rasgele sayılar üretilmiştir. Öncelikle bir ön çalışma yapılarak gruplardaki örneklem büyüklükleri belirlenmeye çalışılmıştır. Bu amaçla, meta analizine dahil edilen çalışma sayısı 3 olduğu durum göz önünde bulundurularak, 7 farklı tedavi etkinlik oranı ve 6 farklı

örneklem büyüklüğü (10, 30, 50, 100, 200 ve 500) için toplam 84 (7x6x2) farklı kombinasyona ait veriler üretilmiştir. Excel programında yazılan bir makro yardımıyla üretilen verilerin Cochran Q heterojenlik testi, testin anlamlılık değeri ve heterojenlik ölçümleri hesaplatılmıştır (61).

Üç farklı çalışmanın meta analizine dahil edildiği durumlar için yapılan simülasyon sonuçlarına göre, grup sayısındaki artışın gruplardaki denek sayıları ile ilişkisini daha iyi ortaya koyabilmek amacıyla, örneklem büyüklüklerinin 50 ve 500 olmasına karar verilmiş ve meta analizine giren çalışma sayısı 5, 10 ve 20 olduğu durumlar için de veriler üretilerek analizler gerçekleştirilmiştir. Tüm işlemler 1000 kez tekrar edilmiştir.

Kanıtı dayalı tıp kapsamında randomize kontrollü çalışmaların meta analizinde grupların tedavi etkinliği bakımından karşılaştırmalarında yaygın kullanımı nedeniyle π değerleri bu çalışmada tedavi etkinlik oranı olarak adlandırılmıştır (62).

3.2. Meta Analizine Dahil Edilen Çalışmaların Heterojen Olduğu Durumlar

3.2.1. Simülasyon Çalışması

Tedavi ve kontrol grupları için tedavi etkinlik oranının her iki grupta birbirinden farklı (heterojen) olduğu bazı kombinasyonlar göz önünde bulundurularak MINITAB 15.0 paket programının demo versiyonunda örneklem büyüklükleri eşit olmak üzere binomiyal dağılım gösteren rasgele sayılar üretilmiştir. Meta analizine dahil edilen çalışma sayısı 3, 5, 10 ve 20 olduğu durumlar için her bir grupta örneklem büyüklüğü eşit (50 ve 500), tedavi ve kontrol grupları için tedavi etkinlik oranları eşit olmayacak şekilde 7 farklı tedavi etkinlik oranları arasındaki fark (0.10; 0.20; 0.30; 0.40; 0.60; 0.80; 0.90) için heterojen yapıda veriler üretilmiştir. Excel programında yazılan bir makro yardımıyla üretilen verilerin Cochran Q heterojenlik testi, testin anlamlılık değeri ve heterojenlik ölçümleri hesaplatılmıştır. Tüm işlemler 1000 kez tekrar edilmiştir. Grafikler STATISTICA 6.0 paket programı ve SPSS 19.0 istatistik paket programının demo versiyonundan elde edilmiştir (63, 64).

4. BULGULAR

4.1. Meta Analizine Dahil Edilen Çalışmaların Homojen Olduğu Durumlarda Heterojenlik Testi ve Heterojenlik Ölçümlerine Ait Bulgular

Meta analizine dahil edilen çalışmaların homojen olduğu durumlarda, çalışma sayısı 3, 5, 10 ve 20 için Cochran Q heterojenlik testi ve p değeri ile heterojenlik ölçümleri ve % 95 güven aralıkları hesaplanarak tablolar ve grafikler halinde sunulmuştur (Çizelge 4.1-4.4, Şekil 4.1-4.6).

Cochran Q'nun değeri, tedavi etkinlik oranı ve örneklem büyüklüğü ne olursa olsun, çalışma sayısı arttıkça sistematik olarak artış göstermektedir. n= 500'de Q değeri n=50'ye nazaran daha fazladır (Çizelge 4.1-4.4, Şekil 4.1).

Heterojenlik ölçümlerinden I^2 değeri incelendiğinde denek sayısının artışına bağlı olarak (n=50'den n=500'e çıkarıldığında) heterojenlik değeri artarken güven aralığı da daralmaktadır. Meta analizine dahil edilen çalışma sayıları bakımından incelendiğinde ise, özellikle küçük örneklerde çalışma sayısı artırıldıkça heterojenlik değerinde bir düşüş gözlenmektedir (Çizelge 4.1-4.4, Şekil 4.2).

Diğer bir heterojenlik ölçüsü olan H incelendiğinde, çalışma sayısı arttıkça ölçüm değeri azalmakta, güven aralıkları da daralmaktadır. H'nin değişkenliği çalışma sayısı az olduğunda ($k < 10$) büyüktür, çalışma sayısı artırıldığında ise yavaşça azalmaktadır. Bu durum örneklem büyüklüğü 500 iken daha bariz gözlenmektedir. Örneklem büyüklüğü küçük olduğunda çalışma sayısı ne olursa olsun, tedavi etkinlik oranlarının 0.05 ve 0.95 olduğu durumlarda heterojenlik ölçümü en küçük değerine sahiptir (Çizelge 4.1-4.4, Şekil 4.3). Yani, bu tedavi etkinlik oranlarında çalışmalar arasındaki homojenlik çok daha kesin bir şekilde ortaya konmaktadır. Heterojenlik ölçümü olarak H kullanılacaksa yanlılığı minimize edebilmek için mutlaka örneklem büyüklüğünü 500, çalışma sayısını da en az 10 olarak belirlemek gerekmektedir. H için söylenenler H_M için de geçerlidir (Çizelge 4.1-4.4, Şekil 4.4). Ayrıca diğer bir heterojenlik ölçümü olan R içinde sonuçlar ve grafikler benzer bulunmuştur (Çizelge 4.1-4.4, Şekil 4.6).

Çalışmalar arası varyans olarak adlandırılan heterojenlik ölçümü τ^2 incelendiğinde, çalışma sayısı ve tedavi etkinlik oranları ne olursa olsun, örneklem sayısı artırıldığında hem τ^2 değeri azalmakta hem de güven aralığı daralmaktadır.

Örneklem büyüklüğü 50 iken, çalışma sayısı artırıldıkça hem τ^2 değeri azalmakta hem de güven aralığı daralmaktadır. Örneklem büyüklüğü 500 iken, çalışma sayısı arttıkça sadece güven aralığında bir daralma gözlenmektedir. Birleştirilmiş çalışmalardaki toplam örneklem büyüklüğü 3000'den sonra τ^2 değeri değişmemektedir. Tedavi etkinlik oranları bakımından incelendiğinde ise, 0.50'ye kadar bir düşüş, 0.50'den sonra artış gözlenmektedir. En küçük τ^2 değeri, tedavi etkinlik oranının 0.50 olduğu durumda, en büyük τ^2 değeri ise tedavi etkinlik oranının 0.05 ve 0.95 olduğu durumda gerçekleşmektedir. Ayrıca güven aralıkları incelendiğinde de en geniş güven aralığı 0.05 ve 0.95 tedavi etkinlik oranı değerlerinde gözlenmektedir (Çizelge 4.1-4.4, Şekil 4.5).

Çizelge 4.1: k=3 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları

k	$\pi_T = \pi_C$	$n_T = n_C$	N	Q (p)	I^2 [%95 CI]	H [%95 CI]	H_M [%95 CI]	τ^2 [%95 CI]	R [%95 CI]
3	0.05	50	300	1.61 (0.53)	10.59 [0.00-85.60]	1.08 [1.00-5.11]	0.44 [0.30-2.14]	0.24 [0.00-17.63]	1.08 [0.27-5.10]
		500	3000	1.96 (0.50)	14.27 [0.68-84.87]	1.13 [1.00-5.02]	0.59 [0.41-2.32]	0.03 [0.00-1.25]	1.13 [0.32-5.01]
	0.10	50	300	1.91 (0.50)	13.96 [0.23-85.57]	1.12 [1.00-4.96]	0.56 [0.38-2.15]	0.18 [0.00-7.68]	1.11 [0.32-4.94]
		500	3000	1.92 (0.51)	13.84 [0.47-85.00]	1.13 [1.00-4.98]	0.57 [0.41-2.21]	0.02 [0.00-0.64]	1.13 [0.32-4.98]
	0.20	50	300	1.89 (0.51)	13.68 [0.65-84.51]	1.12 [1.00-5.03]	0.57 [0.39-2.28]	0.09 [0.00-3.85]	1.12 [0.32-5.02]
		500	3000	2.06 (0.49)	15.66 [0.71-84.30]	1.14 [1.01-4.94]	0.62 [0.41-2.28]	0.01 [0.00-0.36]	1.14 [0.34-4.94]
	0.50	50	300	2.07 (0.49)	15.41 [0.83-85.17]	1.15 [1.01-4.94]	0.62 [0.42-2.23]	0.06 [0.00-2.36]	1.15 [0.34-4.94]
		500	3000	2.04 (0.50)	15.12 [0.87-85.20]	1.14 [1.01-4.95]	0.62 [0.43-2.26]	0.01 [0.00-0.23]	1.14 [0.34-4.95]
	0.80	50	300	1.99 (0.49)	14.69 [0.38-85.46]	1.14 [1.00-4.96]	0.59 [0.42-2.23]	0.10 [0.00-3.85]	1.13 [0.33-4.95]
		500	3000	1.91 (0.51)	13.86 [0.57-83.55]	1.13 [1.00-5.03]	0.58 [0.39-2.29]	0.01 [0.00-0.36]	1.13 [0.32-5.03]
	0.90	50	300	1.95 (0.49)	14.81 [0.21-85.72]	1.13 [1.00-4.94]	0.57 [0.38-2.21]	0.19 [0.00-7.80]	1.12 [0.32-4.93]
		500	3000	2.01 (0.49)	14.97 [0.66-84.81]	1.14 [1.00-4.93]	0.60 [0.43-2.27]	0.02 [0.00-0.63]	1.14 [0.34-4.93]
	0.95	50	300	1.55 (0.54)	9.42 [0.01-85.95]	1.07 [1.00-5.17]	0.41 [0.28-2.16]	0.22 [0.00-17.87]	1.07 [0.26-5.16]
		500	3000	1.98 (0.50)	14.43 [0.79-84.71]	1.14 [1.01-4.99]	0.60 [0.41-2.25]	0.03 [0.00-1.23]	1.14 [0.33-4.98]

Çizelge 4. 2: k=5 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları

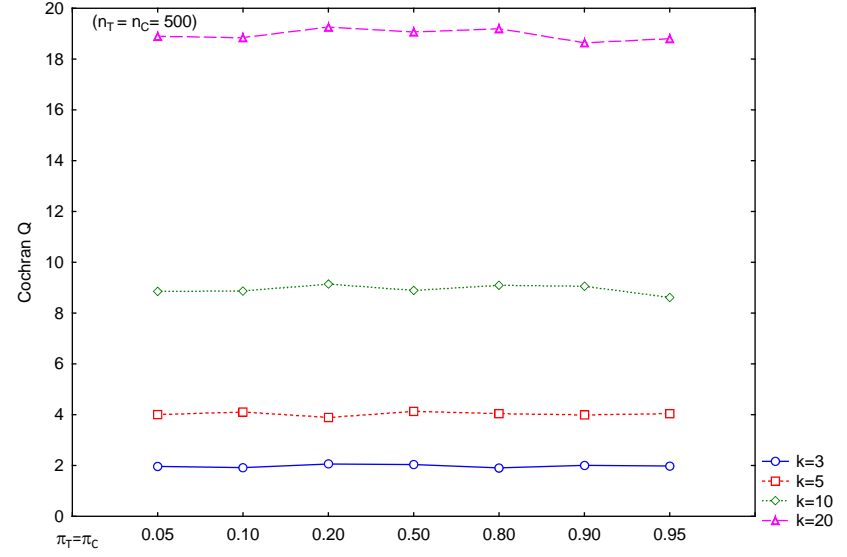
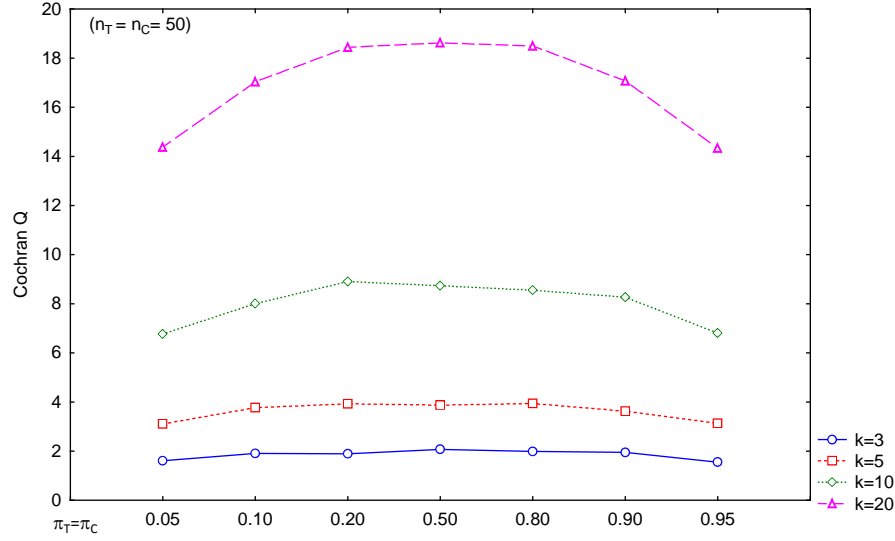
k	$\pi_T = \pi_C$	$n_T = n_C$	N	Q (p)	I^2 [%95 CI]	H [%95 CI]	H_M [%95 CI]	τ^2 [%95 CI]	R [%95 CI]
5	0.05	50	500	3.10 (0.58)	6.01 [0.00-62.40]	1.04 [1.00-2.25]	0.30 [0.29-0.95]	0.11 [0.00-2.72]	1.04 [0.49-2.24]
		500	5000	4.01 (0.50)	14.04 [0.57-63.36]	1.12 [1.00-2.26]	0.53 [0.44-1.26]	0.03 [0.00-0.24]	1.11 [0.56-2.26]
	0.10	50	500	3.77 (0.50)	11.52 [0.11-65.48]	1.09 [1.00-2.24]	0.45 [0.38-1.12]	0.11 [0.00-1.44]	1.08 [0.53-2.23]
		500	5000	4.10 (0.49)	14.35 [0.52-64.71]	1.12 [1.00-2.26]	0.54 [0.44-1.25]	0.01 [0.00-0.13]	1.12 [0.57-2.26]
	0.20	50	500	3.92 (0.50)	12.76 [0.37-65.18]	1.10 [1.00-2.25]	0.50 [0.41-1.20]	0.07 [0.00-0.75]	1.10 [0.55-2.25]
		500	5000	3.90 (0.51)	12.75 [0.71-63.77]	1.10 [1.00-2.26]	0.51 [0.43-1.25]	0.01 [0.00-0.07]	1.10 [0.55-2.26]
	0.50	50	500	3.88 (0.51)	12.54 [0.47-64.52]	1.10 [1.00-2.25]	0.49 [0.41-1.18]	0.04 [0.00-0.45]	1.10 [0.55-2.25]
		500	5000	4.13 (0.49)	14.13 [0.82-64.70]	1.12 [1.01-2.28]	0.55 [0.45-1.29]	0.00 [0.00-0.05]	1.12 [0.56-2.28]
	0.80	50	500	3.94 (0.50)	13.08 [0.48-64.87]	1.11 [1.00-2.26]	0.51 [0.43-1.24]	0.07 [0.00-0.74]	1.10 [0.55-2.25]
		500	5000	4.04 (0.49)	13.51 [0.62-65.57]	1.11 [1.00-2.26]	0.52 [0.43-1.25]	0.01 [0.00-0.07]	1.11 [0.56-2.26]
	0.90	50	500	3.63 (0.52)	10.33 [0.12-64.48]	1.08 [1.00-2.23]	0.42 [0.37-1.09]	0.10 [0.00-1.39]	1.07 [0.53-2.22]
		500	5000	3.99 (0.50)	13.55 [0.49-64.83]	1.11 [1.00-2.26]	0.52 [0.42-1.23]	0.01 [0.00-0.13]	1.11 [0.56-2.26]
	0.95	50	500	3.13 (0.57)	6.79 [0.01-61.51]	1.05 [1.00-2.24]	0.32 [0.310-0.97]	0.13 [0.00-2.69]	1.04 [0.49-2.23]
		500	5000	4.05 (0.49)	14.14 [0.50-64.47]	1.11 [1.00-2.26]	0.53 [0.42-1.23]	0.02 [0.00-0.24]	1.11 [0.56-2.26]

Çizelge 4.3: k=10 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları

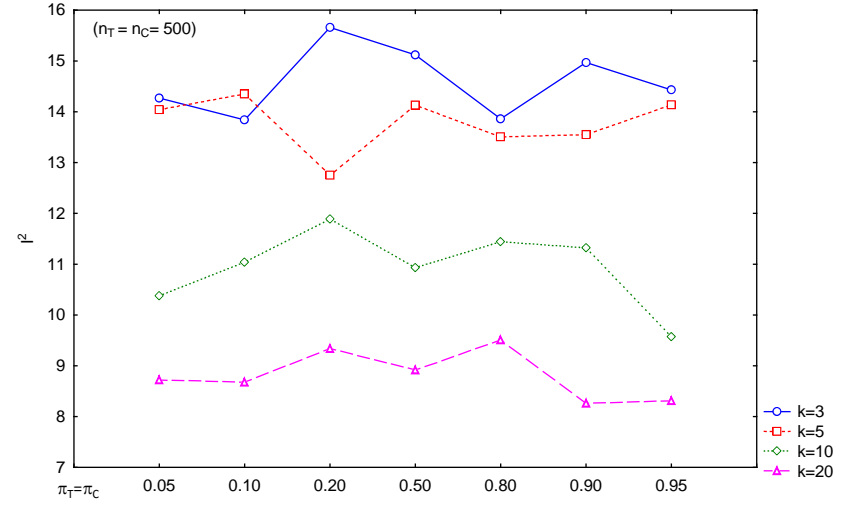
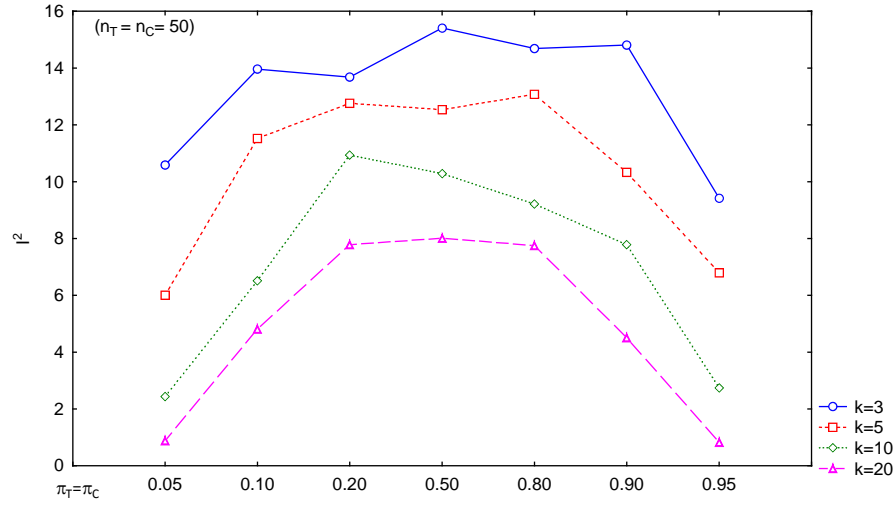
k	$\pi_T = \pi_C$	$n_T = n_C$	N	Q (p)	I^2 [%95 CI]	H [%95 CI]	H_M [%95 CI]	τ^2 [%95 CI]	R [%95 CI]
10	0.05	50	1000	6.77 (0.66)	2.43 [0.00-42.99]	1.01 [1.00-1.63]	0.17 [0.27-0.57]	0.03 [0.00-0.99]	1.01 [0.63-1.63]
		500	10000	8.86 (0.50)	10.38 [0.33-51.03]	1.07 [1.00-1.67]	0.41 [0.39-0.81]	0.01 [0.00-0.11]	1.07 [0.69-1.67]
	0.10	50	1000	8.00 (0.56)	6.52 [0.01-48.93]	1.04 [1.00-1.64]	0.30 [0.33-0.70]	0.05 [0.00-0.60]	1.04 [0.66-1.64]
		500	10000	8.97 (0.51)	11.04 [0.22-50.30]	1.08 [1.00-1.67]	0.42 [0.40-0.84]	0.01 [0.00-0.06]	1.08 [0.70-1.67]
	0.20	50	1000	8.90 (0.50)	10.94 [0.27-50.93]	1.08 [1.00-1.68]	0.42 [0.41-0.85]	0.05 [0.00-0.35]	1.08 [0.69-1.67]
		500	10000	9.14 (0.49)	11.89 [0.38-51.34]	1.08 [1.00-1.68]	0.44 [0.41-0.84]	0.00 [0.00-0.03]	1.08 [0.71-1.68]
	0.50	50	1000	8.74 (0.52)	10.28 [0.42-49.92]	1.07 [1.00-1.67]	0.41 [0.42-0.87]	0.03 [0.00-0.21]	1.07 [0.69-1.67]
		500	10000	8.89 (0.51)	10.93 [0.35-50.67]	1.08 [1.00-1.68]	0.42 [0.41-0.85]	0.00 [0.00-0.02]	1.08 [0.70-1.68]
	0.80	50	1000	8.55 (0.52)	9.26 [0.18-49.98]	1.06 [1.00-1.66]	0.37 [0.36-0.77]	0.04 [0.00-0.34]	1.06 [0.68-1.66]
		500	10000	9.09 (0.49)	11.44 [0.40-51.67]	1.08 [1.00-1.68]	0.44 [0.41-0.86]	0.00 [0.00-0.03]	1.08 [0.70-1.68]
	0.90	50	1000	8.27 (0.54)	7.79 [0.01-49.56]	1.05 [1.00-1.65]	0.33 [0.34-0.71]	0.06 [0.00-0.62]	1.04 [0.67-1.64]
		500	10000	9.06 (0.49)	11.32 [0.26-51.46]	1.08 [1.00-1.67]	0.42 [0.40-0.83]	0.01 [0.00-0.06]	1.08 [0.70-1.67]
	0.95	50	1000	6.81 (0.65)	2.743 [0.00-43.03]	1.02 [1.00-1.63]	0.17 [0.26-0.56]	0.04 [0.00-0.98]	1.01 [0.63-1.63]
		500	10000	8.61 (0.52)	9.57 [0.24-96.98]	1.07 [1.00-1.66]	0.39 [0.37-0.78]	0.01 [0.00-0.11]	1.07 [0.69-1.66]

Çizelge 4.4: k=20 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları

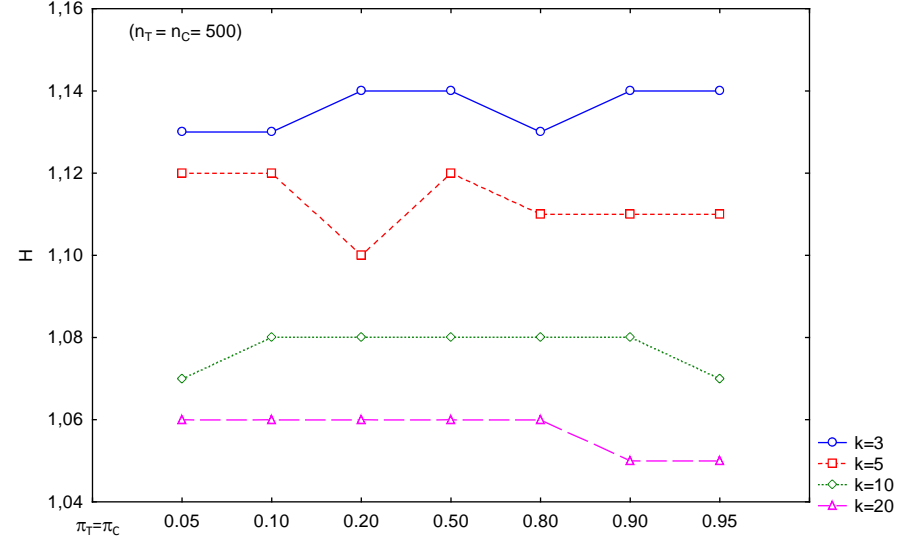
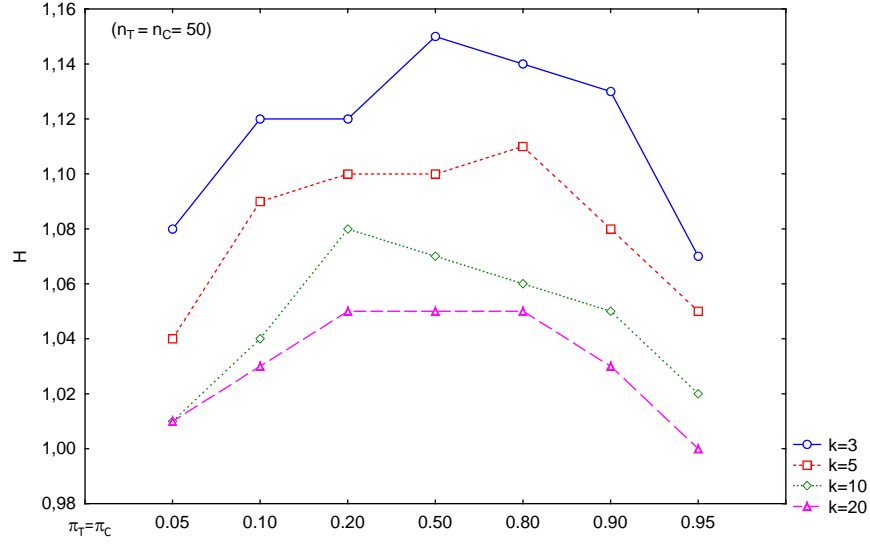
k	$\pi_T = \pi_C$	$n_T = n_C$	N	Q (p)	I^2 [%95 CI]	H [%95 CI]	H_M [%95 CI]	τ^2 [%95 CI]	R [%95 CI]
20	0.05	50	2000	14.38 (0.73)	0.88 [0.00-27.75]	1.01 [1.00-1.93]	0.03 [0.23-0.38]	0.01 [0.00-0.46]	1.00 [0.73-1.83]
		500	20000	18.90 (0.50)	8.72 [0.26-40.17]	1.06 [1.00-1.43]	0.21 [0.37-0.63]	0.01 [0.00-0.07]	1.06 [0.78-1.42]
	0.10	50	2000	17.04 (0.59)	4.81 [0.01-36.15]	1.03 [1.00-1.04]	0.13 [0.31-0.52]	0.03 [0.00-0.33]	1.03 [0.75-1.40]
		500	20000	18.84 (0.51)	8.68 [0.34-39.98]	1.06 [1.00-1.43]	0.21 [0.36-0.62]	0.01 [0.00-0.04]	1.06 [0.78-1.43]
	0.20	50	2000	18.44 (0.52)	7.78 [0.17-39.24]	1.05 [1.00-1.42]	0.19 [0.35-0.60]	0.03 [0.00-0.20]	1.05 [0.78-1.42]
		500	20000	19.26 (0.49)	9.34 [0.33-41.33]	1.06 [1.00-1.43]	0.23 [0.37-0.63]	0.00 [0.00-0.02]	1.06 [0.79-1.43]
	0.50	50	2000	18.62 (0.52)	8.01 [0.26-39.71]	1.05 [1.00-1.42]	0.20 [0.36-0.62]	0.02 [0.00-0.13]	1.05 [0.78-1.42]
		500	20000	19.07 (0.50)	8.92 [0.39-40.69]	1.06 [1.00-1.43]	0.22 [0.37-0.63]	0.00 [0.00-0.01]	1.06 [0.78-1.43]
	0.80	50	2000	18.50 (0.52)	7.75 [0.08-39.56]	1.05 [1.00-1.42]	0.20 [0.34-0.58]	0.03 [0.00-0.20]	1.05 [0.78-1.41]
		500	20000	19.20 (0.49)	9.51 [0.21-40.97]	1.06 [1.00-1.43]	0.23 [0.38-0.64]	0.00 [0.00-0.02]	1.06 [0.79-1.43]
	0.90	50	2000	17.07 (0.59)	4.51 [0.00-36.32]	1.03 [1.00-1.40]	0.13 [0.29-0.49]	0.03 [0.00-0.33]	1.02 [0.75-1.39]
		500	20000	18.64 (0.52)	8.26 [0.25-39.67]	1.05 [1.00-1.42]	0.20 [0.37-0.63]	0.01 [0.00-0.04]	1.05 [0.78-1.42]
	0.95	50	2000	14.34 (0.74)	0.83 [0.00-27.66]	1.00 [1.00-1.39]	0.03 [0.22-0.37]	0.01 [0.00-0.46]	1.00 [0.73-1.38]
		500	20000	18.81 (0.51)	8.31 [0.19-40.42]	1.05 [1.00-1.42]	0.21 [0.36-0.62]	0.01 [0.00-0.07]	1.05 [0.78-1.42]



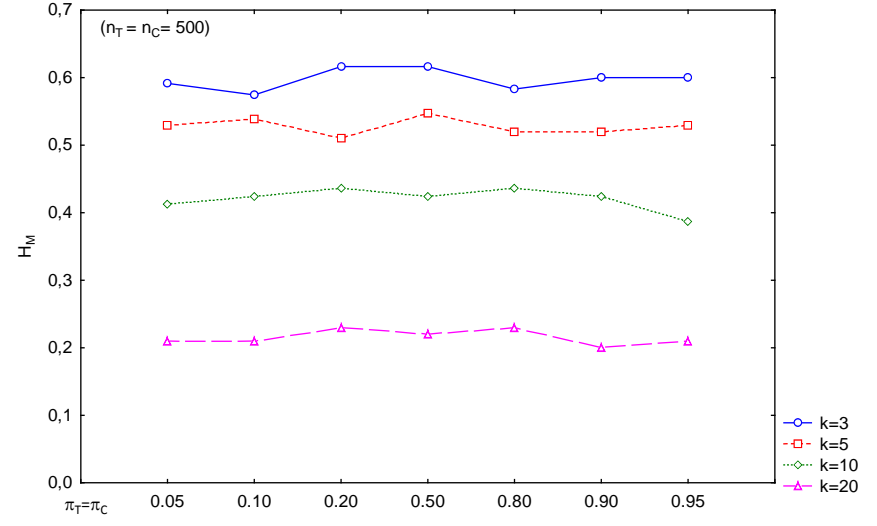
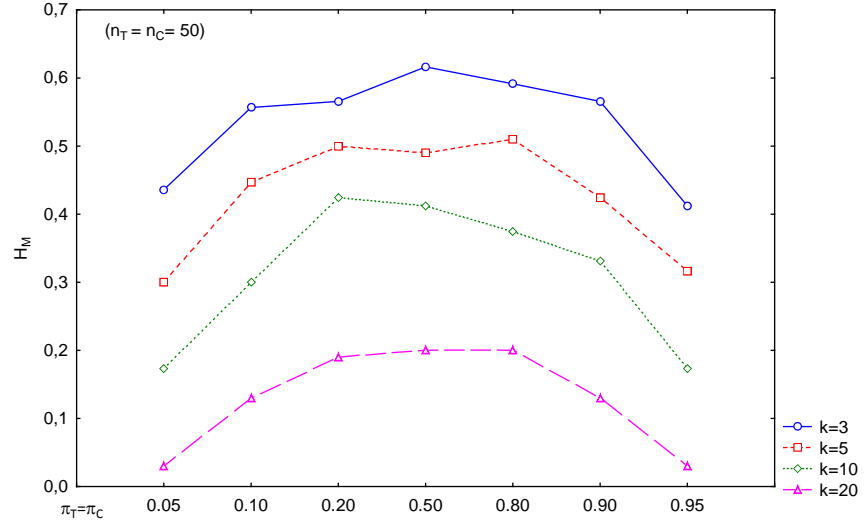
Şekil 4.1: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen Cochran Q



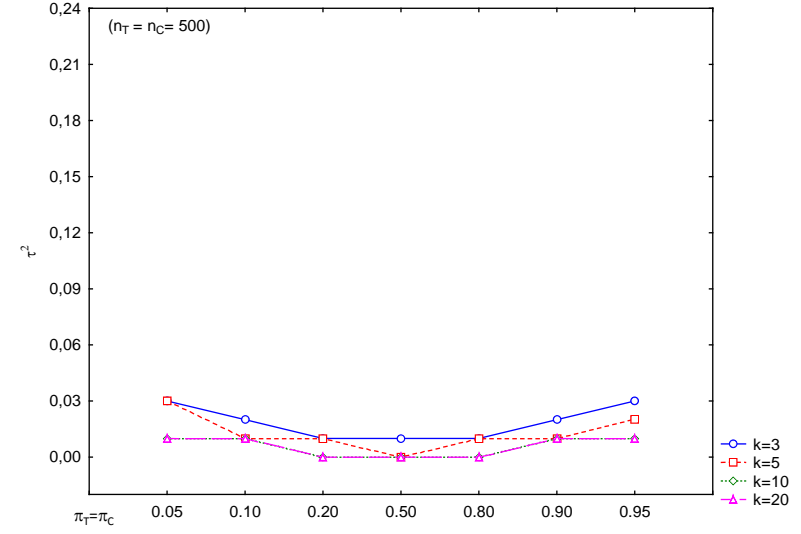
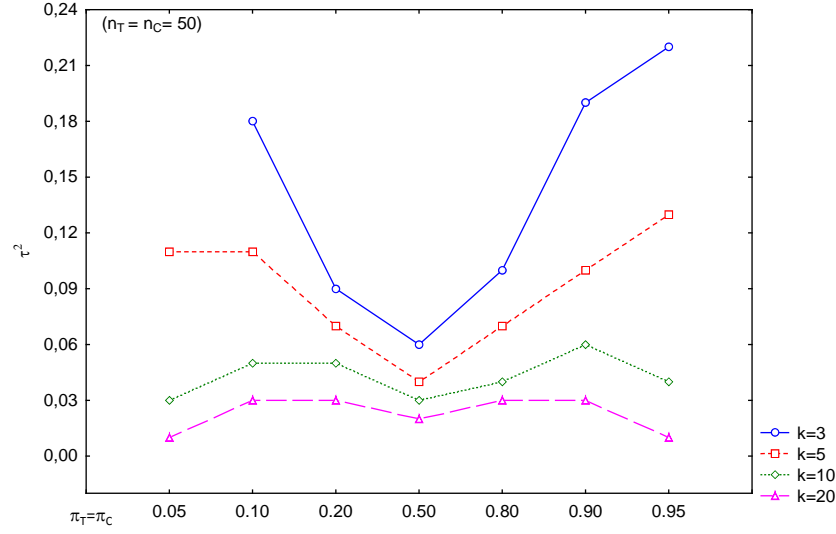
Şekil 4.2: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen I²



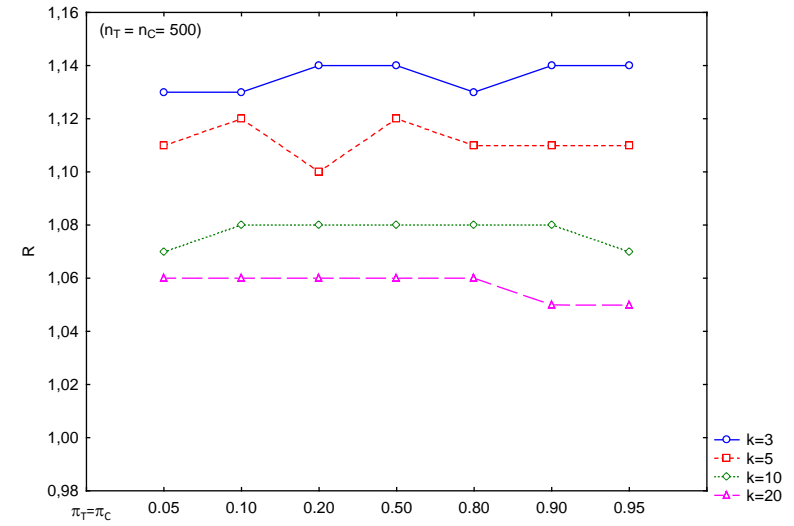
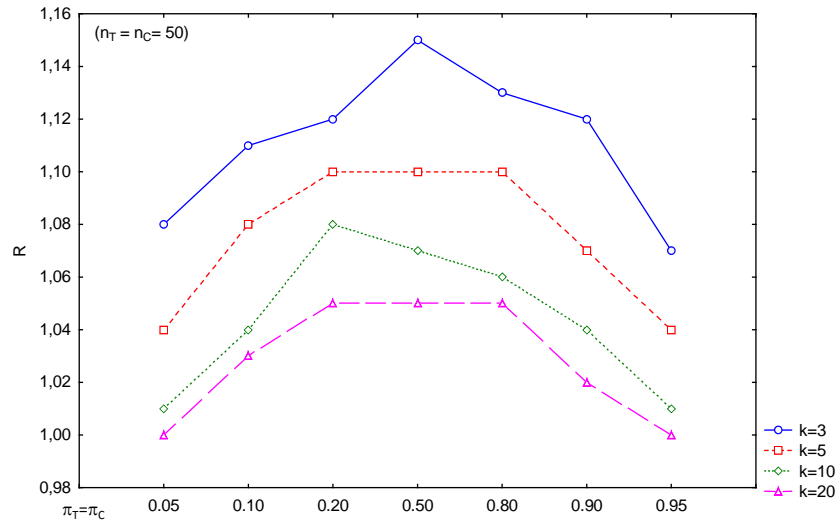
Şekil 4.3: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen H



Şekil 4.4: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen H_M



Şekil 4.5: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen τ^2



Şekil 4.6: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen R

4.2. Meta Analizine Dahil Edilen Çalışmaların Heterojen Olduğu Durumlarda Heterojenlik Testi ve Heterojenlik Ölçümlerine Ait Bulgular

Meta analizine dahil edilen çalışmaların heterojen olduğu durumlarda, çalışma sayısı 3, 5, 10 ve 20 için Cochran Q heterojenlik testi ve p değeri ile heterojenlik ölçümleri ve % 95 güven aralıkları hesaplanarak tablolar ve grafikler halinde sunulmuştur (Çizelge 4.5-4.8, Şekil 4.7-4.12).

Cochran Q değeri incelendiğinde, örneklem büyüklüğü ne olursa olsun çalışma sayısı arttıkça Q değerinin de arttığı gözlenmektedir. $n_T=n_C=500$ 'e ait Q değerleri $n_T=n_C=50$ 'ye nazaran daha fazladır. Ayrıca π değerleri bakımından incelediğimizde ise çalışma sayısı ve örneklem sayısı ne olursa olsun, tedavi ve kontrol gruplarının tedavi etkinlik oranları arasındaki farklılık arttıkça Q değerinde de bir artış gözlenmektedir. Bu artışlar incelendiğinde de en büyük artışın tedavi etkinlik oranları arasındaki farklılığın 0.80'de ($\pi_T=\pi_C=0.90$) olduğu durumda göze çarpmaktadır (Şekil 4.7). Ayrıca, küçük örneklerde ($n_T=n_C=50$) çalışma sayısı 3, 5 ve 10'da, tedavi etkinlik oranları arasındaki farklılığın 0.10'da ($\pi_T=\pi_C=0.80$) homojen durumla karşılaşılmıştır. Heterojen sonuç beklenirken Q testi sonucunda etki büyüklüklerinin homojen olduğu gözlenmiştir. Denek sayısı 100'e hatta 200'e çıkarıldığında ancak heterojenlik durumu sağlanmıştır (Çizelge 4.5-4.8). Bu durum, küçük örneklerde ve meta analizine daha az sayıda çalışma alındığında ($k \leq 10$), tedavi etkinlik oranları arasındaki farkın 0.10 ($\pi_T=\pi_C=0.80$) olduğu durum, Cochran Q testinin gücünün zayıf olduğu anlamına gelmektedir.

Cochran Q değeri ile heterojenlik ölçümü olan I^2 değerine ait sonuçlar, örneklem sayısı büyük olduğunda ($n_T=n_C=500$) birbiri ile tutarlı iken küçük örneklerde ($n_T=n_C=50$) tutarsız olduğu gözlenmektedir. Meta analizine dahil edilen çalışma sayısı 3, örneklem büyüklüğü 50 ve $\pi_T=\pi_C=0.80$ iken Q, I^2 ve % 95 güven aralığı sonuçlarına göre, meta analizine dahil edilen çalışmaların homojen olduğu sonucuna varılmaktadır. Örneklem sayısını 100'e çıkardığımızda Cochran Q'ya ve I^2 'ye ait % 95 güven aralığına göre homojen yorumu yapılırken, I^2 değeri de % 50.91 ile heterojen olarak yorumlanmaktadır. Örneklem sayısını 200'e çıkardığımızda ise Q değeri ve I^2 'ye ait % 95 güven aralığı sonucunda heterojen yorumu yapılırken, I^2 değeri de % 70.44 değeri ile heterojenliğin varlığını desteklemektedir. $\pi_T=\pi_C=0.80$ 'de ve çalışma sayısı 20 olduğunda ise, Q'ya göre, meta analizine dahil edilen çalışmaların heterojen olduğu yorumu yapılırken I^2 değeri % 43.69 olarak hesaplanmıştır. Buna göre Q ile heterojenlik ölçümü arasında bir tutarsızlık gözlenmektedir (Çizelge 4.5-4.8).

Heterojenlik ölçümüne ait grafikler incelendiğinde ise, örneklem sayısı 500’de çalışma sayısı ve tedavi etkinlik oranları arasındaki farklılık ne olursa olsun I^2 ’ye göre heterojen bir durum olduğu sonucuna varılmaktadır. Ayrıca tedavi etkinlik oranları arasındaki farklılık bakımından incelediğimizde, küçük örneklemelerde çalışma sayısı ne olursa olsun, tedavi ve kontrol gruplarının tedavi etkinlik oranları arasındaki farklılık arttıkça I^2 değerinde de bir artış gözlenirken, örneklem büyüklüğü artırıldığında çalışma sayısından ve tedavi ve kontrol gruplarının tedavi etkinlik oranları arasındaki farklılıklardan etkilenmediği gözlenmektedir. Küçük örneklemelerde ($n_T=n_C=50$), tedavi etkinlik oranları arasındaki farklılık 0.10’da I^2 değeri % 30 ile % 40 arasında bir değer alırken, farklılık 0.20’den sonra I^2 değeri % 70’in üzerine çıkmaktadır. Büyük örneklemelerde ise, tedavi etkinlik oranları arasındaki farklılığın 0.10 olduğu durum da bile I^2 değerinin % 80’in üzerinde bir değer aldığı gözlenmektedir (Şekil 4.8).

I^2 değeri meta analizine dahil edilen çalışma sayısından, örneklem sayısından ve tedavi etkinlik oranları arasındaki farklılıktan etkilenmemekle birlikte $n_T=n_C=500$ ’deki değerler $n_T=n_C=50$ ’deki değerlere nazaran biraz daha fazladır. I^2 değerine ait güven aralıkları incelendiğinde, her bir π değeri için örneklem sayısı ne olursa olsun çalışma sayısı arttıkça güven aralıklarında bir daralma gözlenmektedir. Örneklem sayıları bakımından karşılaştırıldığında ise büyük örneklemelerde küçük örneklemelere nazaran güven aralıkları biraz daha daralmaktadır. Ayrıca, $\pi_T=\pi_C=0.80$ ’de karşılaşılan homojen durumlarda güven aralıkları oldukça geniş heterojen duruma dönüştüğünde ise güven aralıklarının daraldığı söylenebilmektedir (Çizelge 4.5-4.8).

H, H_M ve R heterojenlik ölçümleri incelendiğinde, çalışma sayısı 3 ve 5’de değişkenlik yüksek, çalışma sayısı artırıldığında bir düşüş gözlenmektedir. Yalnız bu düşüşler çok da bariz değişiklikler değildir (Çizelge 4.5-4.8). Örneklem büyüklükleri bakımından incelendiğinde, küçük örneklemelerde heterojenlik ölçümleri küçük iken örneklem sayısı arttığında heterojenlik ölçümleri 3 katı kadar fazla değer almaktadır. Tedavi etkinlik oranları arasındaki farklılıklar bakımından incelendiğinde ise, farklılık arttıkça heterojenlik ölçümlerine ait değerler artmaktadır. En büyük değer, farklılığın 0.80 olduğu durumda gözlenmektedir. 0.90’da ise 0.80’e nazaran bir düşüş olmakla birlikte yine de bu düşüş diğer etkinlik oranları farklılıklarına göre daha fazladır (Şekil 4.9, 4.10 ve 4.12).

τ^2 heterojenlik ölçümü incelendiğinde, çalışma sayısından ve örneklem sayısından etkilenmediği, sadece tedavi etkinlik oranları arasındaki farklılıklardan etkilendiği göze çarpmaktadır. Buna göre, tedavi etkinlik oranları arasındaki farklılık arttıkça τ^2 değeri de,

güven aralığı da artmaktadır ve en geniş güven aralığının $\pi_T - \pi_C = 0.90$ olduğu durumda gerçekleşmektedir (Çizelge 4.5-4.8, Şekil 4.11).

Çizelge 4.5: k=3 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları

k	1	2	3	$n_T=n_C$	N	Q (p)	I^2 [%95 CI]	H [%95 CI]	H_M [%95 CI]	τ^2 [%95 CI]	R [%95 CI]
3	$\pi_T=0.80$ $\pi_C=0.20$	$\pi_T=0.05$ $\pi_C=0.05$	$\pi_T=0.20$ $\pi_C=0.80$	50	300	61.49 (0.00)	96.68 (93.15-98.39)	5.53 (3.86-7.92)	5.43 (3.80-7.79)	12.17 (5.76-25.28)	4.71 (3.29-6.75)
				500	3000	615.59 (0.00)	99.67 (99.54-99.77)	1754 (14.82-20.76)	17.51 (14.80-20.72)	11.18 (7.98-15.67)	15.67 (13.24-18.54)
	$\pi_T=0.50$ $\pi_C=0.10$	$\pi_T=0.10$ $\pi_C=0.10$	$\pi_T=0.10$ $\pi_C=0.50$	50	300	30.80 (0.00)	93.19 (83.29-97.21)	3.90 (2.52-6.05)	3.79 (2.43-5.85)	5.69 (2.16-14.15)	3.77 (2.44-5.85)
				500	3000	318.32 (0.00)	99.37 (99.04-99.59)	12.61 (10.21-15.58)	12.58 (10.17-15.53)	5.50 (3.59-8.40)	12.47 (10.09-15.41)
	$\pi_T=0.50$ $\pi_C=0.20$	$\pi_T=0.20$ $\pi_C=0.20$	$\pi_T=0.20$ $\pi_C=0.50$	50	300	19.88 (0.00)	88.25 (67.35-95.83)	3.10 (1.90-5.08)	2.99 (1.80-4.79)	2.09 (0.66-5.89)	3.08 (1.89-5.04)
				500	3000	188.99 (0.00)	98.93 (98.22-99.35)	9.70 (7.54-12.48)	9.67 (7.50-12.42)	2.06 (1.24-3.42)	9.68 (7.52-12.45)
	$\pi_T=0.50$ $\pi_C=0.30$	$\pi_T=0.50$ $\pi_C=0.50$	$\pi_T=0.30$ $\pi_C=0.50$	50	300	10.21 (0.06)	71.43 (31.85-92.78)	2.18 (1.31-4.10)	2.03 (1.11-3.47)	0.73 (0.15-2.86)	2.17 (1.21-4.10)
				500	3000	83.81 (0.00)	97.50 (95.12-98.72)	6.44 (4.64-8.93)	6.40 (4.59-8.82)	0.70 (0.36-1.35)	6.43 (4.64-8.92)
	$\pi_T=0.50$ $\pi_C=0.40$	$\pi_T=0.80$ $\pi_C=0.80$	$\pi_T=0.40$ $\pi_C=0.50$	50	300	4.09 (0.29)	35.64 (5.65-89.47)	1.41 (1.04-4.36)	1.10 (0.63-2.56)	0.24 (0.02-2.90)	1.40 (0.59-4.34)
				100	600	6.14 (0.18)	50.91 (13.30-90.73)	1.68 (1.11-4.10)	1.47 (0.81-2.88)	0.21 (0.03-1.39)	1.67 (0.82-4.07)
				200	1200	10.38 (0.07)	70.44 (31.71-92.53)	2.18 (1.33-4.15)	2.05 (1.13-3.50)	0.20 (0.04-0.78)	2.16 (1.21-4.11)
				500	3000	21.94 (0.00)	88.77 (69.24-96.00)	3.24 (2.01-5.24)	3.16 (1.91-4.96)	0.19 (0.06-0.51)	3.20 (1.99-5.19)
	$\pi_T=0.90$ $\pi_C=0.10$	$\pi_T=0.90$ $\pi_C=0.90$	$\pi_T=0.10$ $\pi_C=0.90$	50	300	79.02 (0.00)	97.44 (95.00-98.69)	6.28 (4.50-8.75)	6.20 (4.44-8.64)	20.41 (10.23-40.25)	6.12 (4.35-8.54)
				500	3000	862.75 (0.00)	99.77 (99.69-99.83)	20.77 (17.89-24.12)	20.75 (17.87-24.09)	19.34 (14.33-26.08)	20.74 (17.87-24.08)
	$\pi_T=0.95$ $\pi_C=0.05$	$\pi_T=0.95$ $\pi_C=0.95$	$\pi_T=0.05$ $\pi_C=0.95$	50	300	67.30 (0.00)	96.94 (93.79-98.49)	5.78 (4.08-8.20)	5.71 (4.02-8.08)	37.04 (17.74-76.14)	5.54 (3.91-7.86)
				500	3000	807.67 (0.00)	99.75 (99.66-99.82)	20.09 (17.24-23.41)	20.08 (17.22-23.38)	35.07 (25.80-47.66)	20.03 (17.19-23.34)

Çizelge 4.6: k=5 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları

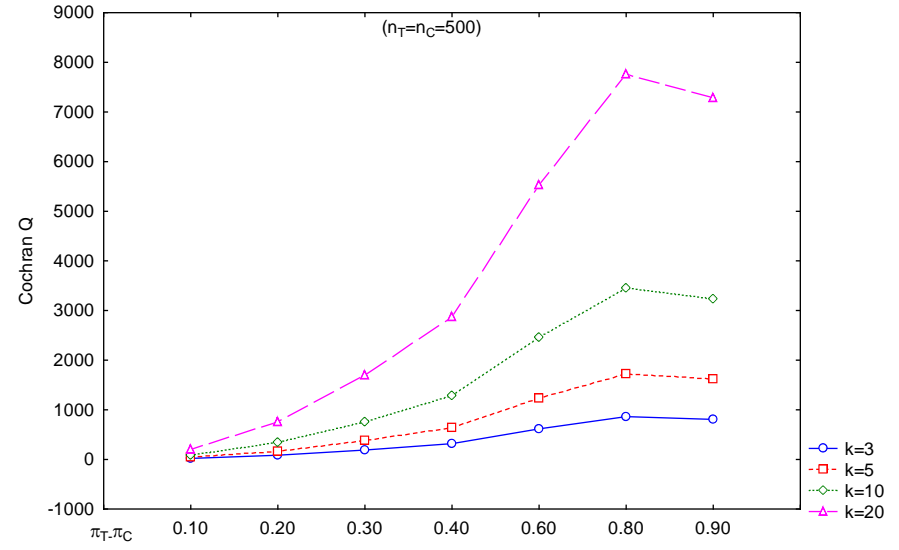
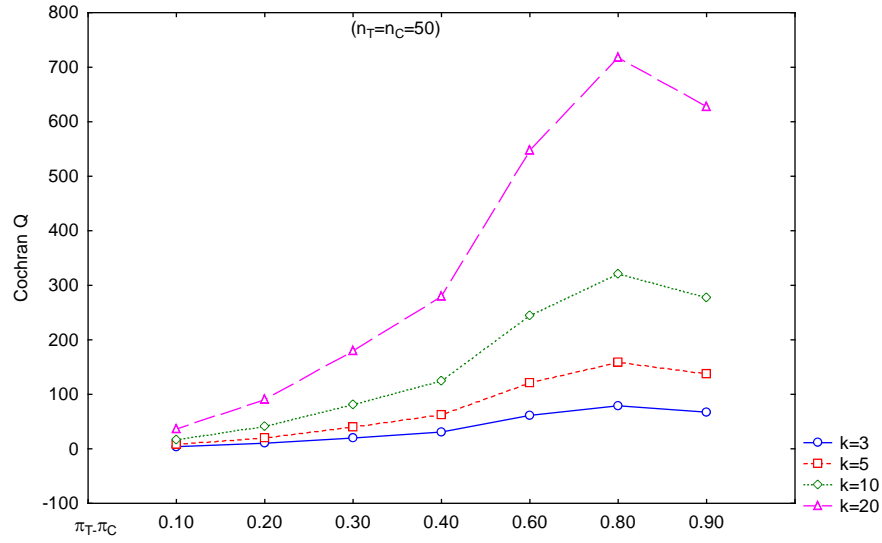
k	1-2	3	4-5	$n_T=n_C$	N	Q (p)	I^2 [%95 CI]	H [%95 CI]	H_M [%95 CI]	τ^2 [%95 CI]	R [%95 CI]
5	$\pi_T=0.80$ $\pi_C=0.20$	$\pi_T=0.05$ $\pi_C=0.05$	$\pi_T=0.20$ $\pi_C=0.80$	50	500	121.49 (0.00)	96.67 (94.41-98.02)	5.505 (4.25-7.13)	5.41 (4.18-7.01)	9.48 (5.54-16.09)	4.69 (3.62-6.08)
				500	5000	1229.44 (0.00)	99.67 (99.59-99.74)	17.53 (15.55-19.76)	17.50 (15.53-19.72)	9.23 (7.268-11.72)	15.71 (13.94-17.71)
	$\pi_T=0.50$ $\pi_C=0.10$	$\pi_T=0.10$ $\pi_C=0.10$	$\pi_T=0.10$ $\pi_C=0.50$	50	500	61.69 (0.00)	93.37 (87.41-96.50)	3.92 (2.85-5.38)	3.80 (2.76-5.20)	5.14 (2.57-9.97)	3.72 (2.71-5.11)
				500	5000	636.82 (0.00)	99.37 (99.15-99.53)	12.61 (10.85-14.66)	12.58 (10.82-14.62)	5.17 (3.82-7.00)	12.46 (10.72-14.48)
	$\pi_T=0.50$ $\pi_C=0.20$	$\pi_T=0.20$ $\pi_C=0.20$	$\pi_T=0.20$ $\pi_C=0.50$	50	500	39.92 (0.00)	89.25 (77.45-94.86)	3.13 (2.19-4.49)	3.00 (2.07-4.25)	2.01 (0.88-4.32)	3.10 (2.17-4.44)
				500	5000	376.24 (0.00)	98.93 (98.46-99.25)	9.69 (8.10-11.60)	9.65 (8.05-11.54)	1.99 (1.38-2.86)	9.66 (8.07-11.56)
	$\pi_T=0.50$ $\pi_C=0.30$	$\pi_T=0.50$ $\pi_C=0.50$	$\pi_T=0.30$ $\pi_C=0.50$	50	500	19.68 (0.02)	75.59 (44.21-90.21)	2.17 (1.43-3.34)	1.98 (1.25-2.94)	0.70 (0.21-1.86)	2.17 (1.41-3.34)
				500	5000	166.21 (0.00)	97.54 (96.05-98.47)	6.43 (5.09-8.13)	6.37 (5.02-8.03)	0.70 (0.43-1.13)	6.43 (5.08-8.12)
	$\pi_T=0.50$ $\pi_C=0.40$	$\pi_T=0.80$ $\pi_C=0.80$	$\pi_T=0.40$ $\pi_C=0.50$	50	500	7.83 (0.24)	37.18 (6.65-78.24)	1.38 (1.05-2.45)	1.02 (0.66-1.75)	0.19 (0.02-0.86)	1.37 (0.79-2.43)
				100	1000	11.78 (0.10)	56.83 (17.62-84.33)	1.67 (1.14-2.73)	1.40 (0.85-2.16)	0.18 (0.03-0.59)	1.66 (1.02-2.71)
				200	2000	20.43 (0.02)	76.06 (45.62-90.34)	2.21 (1.46-3.39)	2.03 (1.29-2.99)	0.18 (0.06-0.48)	2.19 (1.43-3.35)
	$\pi_T=0.90$ $\pi_C=0.10$	$\pi_T=0.90$ $\pi_C=0.90$	$\pi_T=0.10$ $\pi_C=0.90$	50	500	158.96 (0.00)	97.47 (95.93-98.43)	6.30 (4.97-7.99)	6.22 (4.91-7.89)	19.67 (12.03-31.96)	6.03 (4.75-7.64)
				500	5000	1726.43 (0.00)	99.77 (99.71-99.81)	20.78 (18.69-23.10)	20.75 (18.66-23.07)	19.41 (15.69-24.00)	20.73 (18.64-23.05)
	$\pi_T=0.95$ $\pi_C=0.05$	$\pi_T=0.95$ $\pi_C=0.95$	$\pi_T=0.05$ $\pi_C=0.95$	50	500	136.85 (0.00)	97.04 (95.10-98.21)	5.84 (4.55-7.49)	5.76 (4.48-7.38)	35.75 (21.24-59.72)	5.41 (4.21-6.94)
				500	5000	1617.42 (0.00)	99.75 (99.69-99.80)	20.11 (18.04-22.41)	20.08 (18.02-22.38)	34.88 (28.06-43.36)	19.99 (17.94-22.28)

Çizelge 4.7: k=10 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları

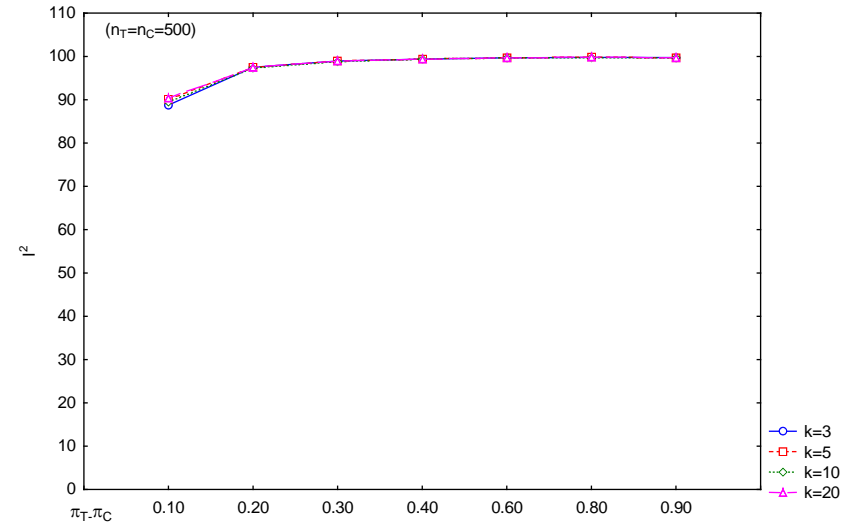
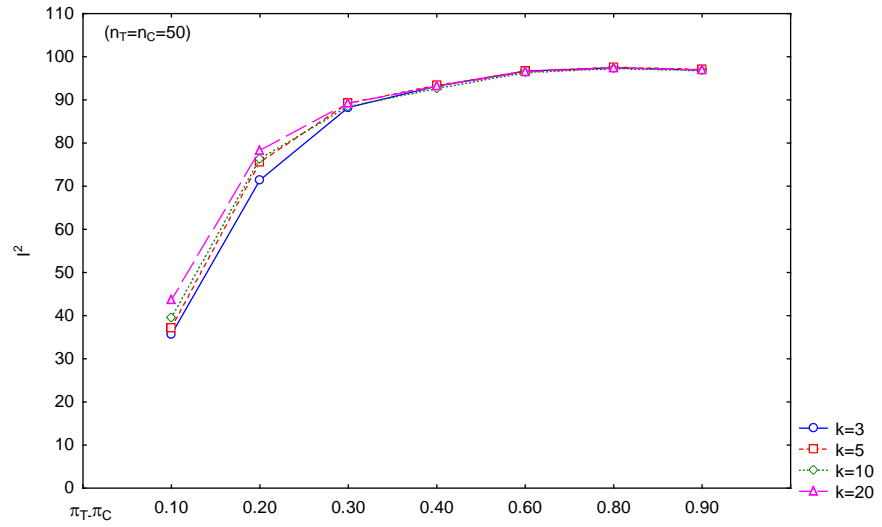
k	1-4	5-6	7-10	$n_T=n_C$	N	Q (p)	I^2 [%95 CI]	H [%95 CI]	H_M [%95 CI]	τ^2 [%95 CI]	R [%95 CI]
10	$\pi_T=0.80$ $\pi_C=0.20$	$\pi_T=0.05$ $\pi_C=0.05$	$\pi_T=0.20$ $\pi_C=0.80$	50	1000	243.96 (0.00)	96.29(94.68-97.42)	5.20 (4.35-6.23)	5.11 (4.27-6.11)	8.20 (5.65-11.89)	4.35 (3.64-5.21)
				500	10000	2459.89 (0.00)	99.63 (99.57-99.69)	16.53 (15.21-17.97)	16.50 (15.18-17.93)	8.07 (6.83-9.54)	14.68 (13.51-15.96)
	$\pi_T=0.50$ $\pi_C=0.10$	$\pi_T=0.10$ $\pi_C=0.10$	$\pi_T=0.10$ $\pi_C=0.50$	50	1000	124.64 (0.00)	92.71 (88.63-95.32)	3.72 (2.98-4.64)	3.58 (2.87-4.46)	4.51 (2.78-7.20)	3.49 (2.80-4.35)
				500	10000	1278.21 (0.00)	99.30 (99.13-99.43)	11.92 (10.73-13.23)	11.86 (10.69-13.18)	4.61 (3.73-5.69)	11.75 (10.58-13.04)
	$\pi_T=0.50$ $\pi_C=0.20$	$\pi_T=0.20$ $\pi_C=0.20$	$\pi_T=0.20$ $\pi_C=0.50$	50	1000	81.00 (0.00)	88.51 (80.90-93.09)	2.99 (2.33-3.84)	2.83 (2.19-3.61)	1.78 (1.00-3.08)	2.95 (2.30-3.79)
				500	10000	755.73 (0.00)	98.80 (98.46-99.07)	9.16 (8.08-10.38)	9.11 (8.03-10.32)	1.77 (1.38-2.28)	9.13 (8.05-10.34)
	$\pi_T=0.50$ $\pi_C=0.30$	$\pi_T=0.50$ $\pi_C=0.50$	$\pi_T=0.30$ $\pi_C=0.50$	50	1000	41.24 (0.00)	76.33 (56.07-87.26)	2.12 (1.57-2.87)	1.89 (1.38-2.51)	0.64 (0.27-1.30)	2.12 (1.57-2.86)
				500	10000	335.84 (0.00)	97.29 (96.24-98.05)	6.10 (5.18-7.18)	6.03 (5.11-7.08)	0.63 (0.45-0.87)	6.10 (5.18-7.17)
	$\pi_T=0.50$ $\pi_C=0.40$	$\pi_T=0.80$ $\pi_C=0.80$	$\pi_T=0.40$ $\pi_C=0.50$	50	1000	16.77 (0.15)	39.85 (8.09-70.43)	1.35 (1.06-1.95)	0.94 (0.66-1.33)	0.16 (0.02-0.51)	1.34 (0.93-1.94)
				100	2000	24.94 (0.04)	58.57 (25.67-79.29)	1.64 (1.20-2.30)	1.33 (0.92-1.80)	0.16 (0.04-0.39)	1.62 (1.16-2.28)
				200	4000	40.63 (0.00)	75.71 (54.89-86.99)	2.10 (1.56-2.84)	1.87 (1.36-2.49)	0.16 (0.07-0.32)	2.08 (1.54-2.81)
				500	10000	89.39 (0.00)	89.46 (82.67-93.58)	3.13 (2.46-4.00)	2.99 (2.33-3.79)	0.16 (0.09-0.27)	3.10 (2.43-3.95)
	$\pi_T=0.90$ $\pi_C=0.10$	$\pi_T=0.90$ $\pi_C=0.90$	$\pi_T=0.10$ $\pi_C=0.90$	50	1000	320.49 (0.00)	97.19 (96.09-97.98)	5.97 (5.06-7.04)	5.88 (4.99-6.94)	17.18 (12.21-24.10)	5.66 (4.80-6.67)
				500	10000	3452.54 (0.00)	99.74 (99.70-99.78)	19.59 (18.19-21.09)	19.56 (18.17-21.06)	17.17 (14.80-19.91)	19.53 (18.14-21.02)
	$\pi_T=0.95$ $\pi_C=0.05$	$\pi_T=0.95$ $\pi_C=0.95$	$\pi_T=0.05$ $\pi_C=0.95$	50	1000	277.34 (0.00)	96.74 (95.38-97.69)	5.55 (4.67-6.59)	5.46 (4.59-6.49)	30.66 (21.38-43.79)	5.00 (4.20-5.94)
				500	10000	3236.91 (0.00)	99.72 (99.68-99.76)	18.96 (17.58-20.45)	18.94 (17.56-20.42)	30.91 (26.56-35.97)	18.82 (17.45-20.29)

Çizelge 4.8: k=20 için heterojenlik testi ve anlamlılık seviyesi, heterojenlik ölçümleri değerleri ve % 95 güven aralıkları

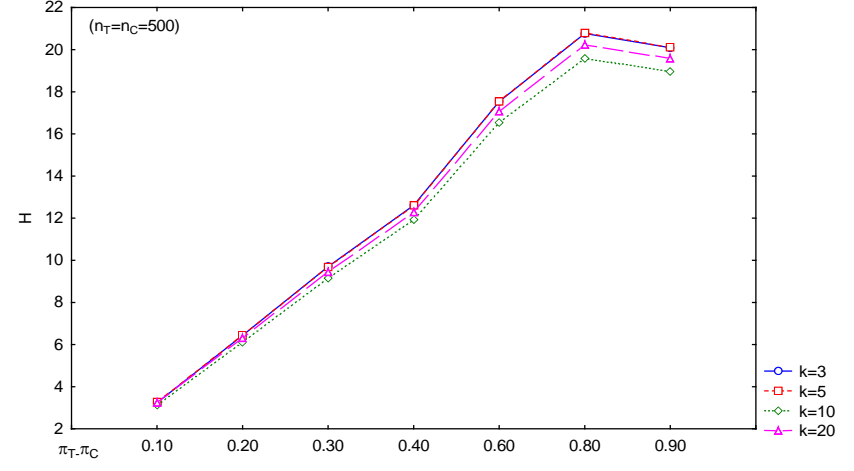
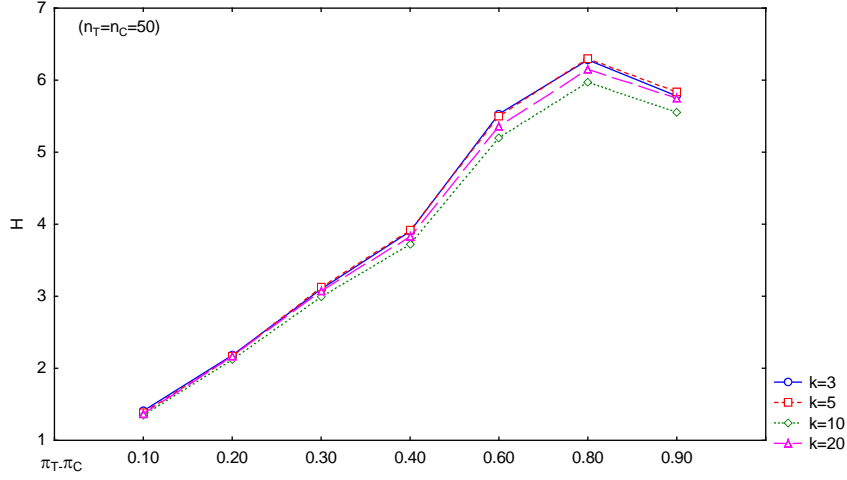
k	1-9	10-11	12-20	$n_T=n_C$	N	Q (p)	I^2 [%95 CI]	H [%95 CI]	H_M [%95 CI]	τ^2 [%95 CI]	R [%95 CI]
20	$\pi_T=0.80$ $\pi_C=0.20$	$\pi_T=0.05$ $\pi_C=0.05$	$\pi_T=0.20$ $\pi_C=0.80$	50	2000	546.89 (0.00)	96.52 (95.56-97.27)	5.36 (4.75-6.06)	5.27 (4.66-5.95)	7.91 (6.14-10.17)	4.76 (4.22-5.38)
				500	20000	5531.66 (0.00)	99.66 (99.62-99.69)	17.06 (16.13-18.05)	17.03 (16.10-18.02)	7.86 (7.03-8.80)	15.87 (15.00-16.78)
	$\pi_T=0.50$ $\pi_C=0.10$	$\pi_T=0.10$ $\pi_C=0.10$	$\pi_T=0.10$ $\pi_C=0.50$	50	2000	279.55 (0.00)	93.17 (90.78-94.95)	3.83 (3.30-4.45)	3.70 (3.19-4.30)	4.56 (3.30-6.28)	3.58 (3.09-4.16)
				500	20000	2873.49 (0.00)	99.34 (99.24-99.43)	12.30 (11.46-13.20)	12.26 (11.42-13.15)	4.72 (4.10-5.44)	12.17 (11.34-13.06)
	$\pi_T=0.50$ $\pi_C=0.20$	$\pi_T=0.20$ $\pi_C=0.20$	$\pi_T=0.20$ $\pi_C=0.50$	50	2000	179.64 (0.00)	89.26 (84.85-92.39)	3.07 (2.59-3.64)	2.90 (2.44-3.44)	1.83 (1.24-2.66)	3.03 (2.56-3.60)
				500	20000	1700.21 (0.00)	98.88 (98.67-99.05)	9.46 (8.69-10.29)	9.40 (8.64-10.23)	1.85 (1.56-2.20)	9.43 (8.67-10.26)
	$\pi_T=0.50$ $\pi_C=0.30$	$\pi_T=0.50$ $\pi_C=0.50$	$\pi_T=0.30$ $\pi_C=0.50$	50	2000	90.11 (0.00)	78.22 (66.79-85.71)	2.17 (1.76-2.67)	1.92 (1.56-2.37)	0.67 (0.38-1.10)	2.16 (1.76-2.66)
				500	20000	755.65 (0.00)	97.47 (96.85-97.97)	6.30 (5.65-7.04)	6.22 (5.57-6.95)	0.67 (0.54-0.85)	6.30 (5.64-7.03)
	$\pi_T=0.50$ $\pi_C=0.40$	$\pi_T=0.80$ $\pi_C=0.80$	$\pi_T=0.40$ $\pi_C=0.50$	50	2000	36.32 (0.06)	43.69 (14.34-66.52)	1.37 (1.10-1.78)	0.91(0.72-1.20)	0.16 (0.04-0.38)	1.36 (1.05-1.77)
				500	20000	200.99 (0.00)	90.38 (86.56-93.11)	3.25 (2.75-3.83)	3.09 (2.62-3.64)	0.16 (0.11-0.23)	3.22 (2.73-3.80)
	$\pi_T=0.90$ $\pi_C=0.10$	$\pi_T=0.90$ $\pi_C=0.90$	$\pi_T=0.10$ $\pi_C=0.90$	50	2000	718.05 (0.00)	97.35 (96.69-97.88)	6.15 (5.50-6.87)	6.07 (5.42-6.78)	18.25 (14.49-22.95)	5.77 (5.16-6.45)
				500	20000	7767.33 (0.00)	99.76 (99.73-99.78)	20.22 (19.24-21.25)	20.19 (19.21-21.22)	18.29 (16.56-20.21)	20.15 (19.17-21.18)
	$\pi_T=0.95$ $\pi_C=0.05$	$\pi_T=0.95$ $\pi_C=0.95$	$\pi_T=0.05$ $\pi_C=0.95$	50	2000	627.58 (0.00)	96.97 (96.17-97.60)	5.75 (5.11-6.46)	5.66 (5.03-6.36)	32.23 (25.30-40.99)	5.10 (4.54-5.74)
				500	20000	7286.49 (0.00)	99.74 (99.71-99.76)	19.58 (18.61-20.60)	19.56 (18.59-20.58)	32.85 (29.67-36.38)	19.41 (18.45-20.43)



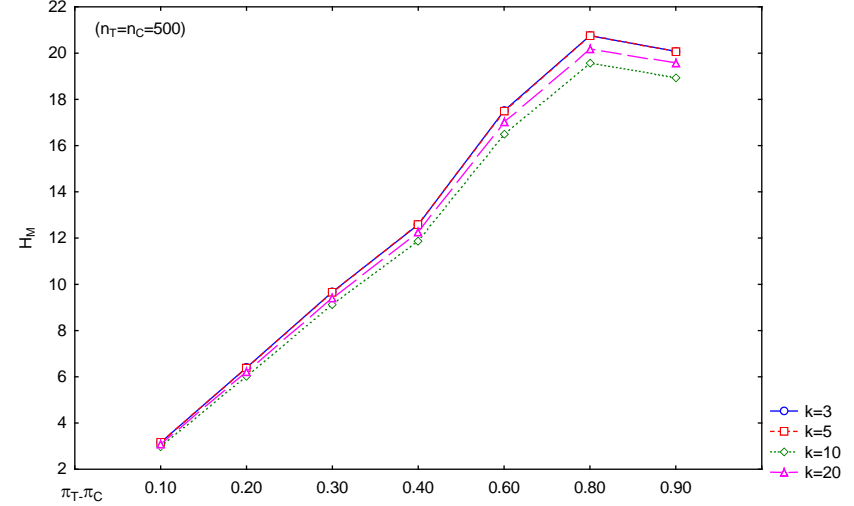
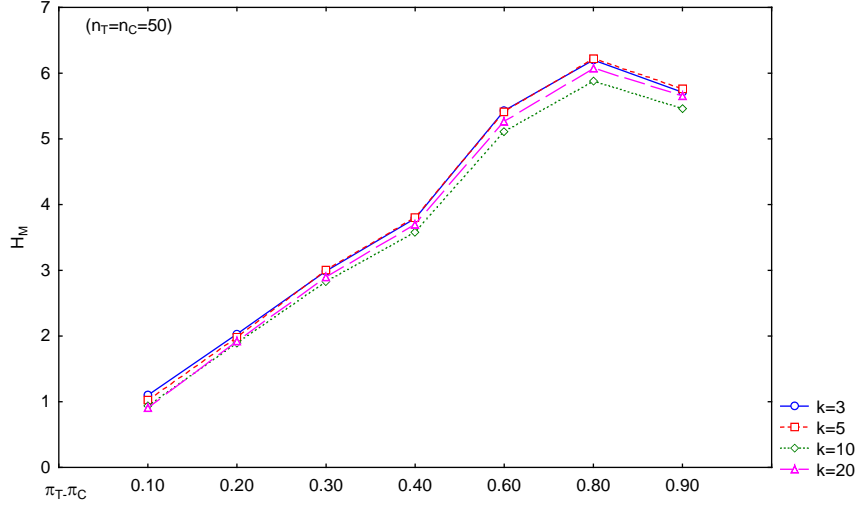
Şekil 4.7: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen Cochran Q



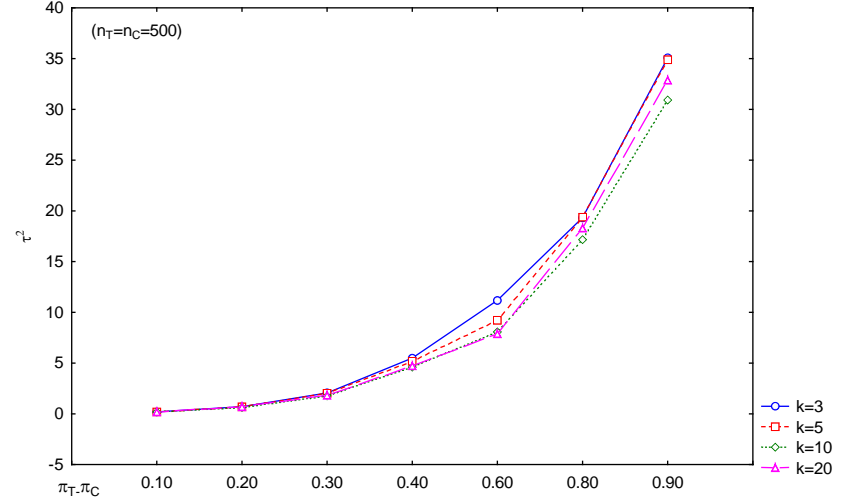
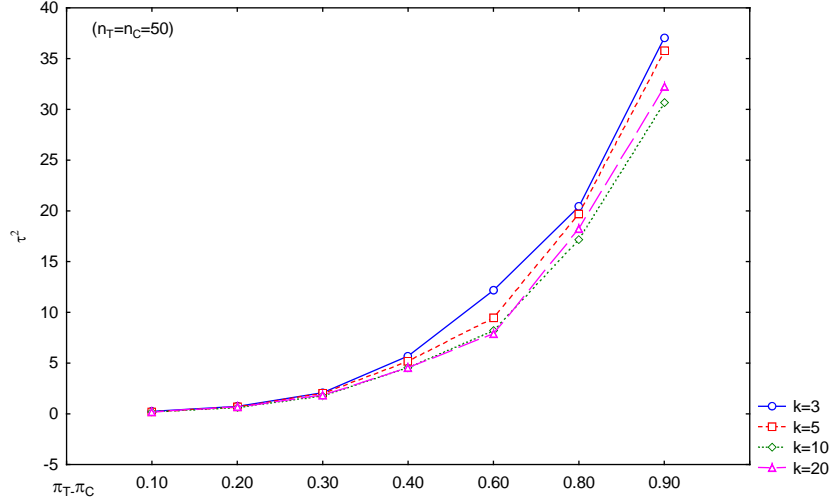
Şekil 4.8: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen I^2



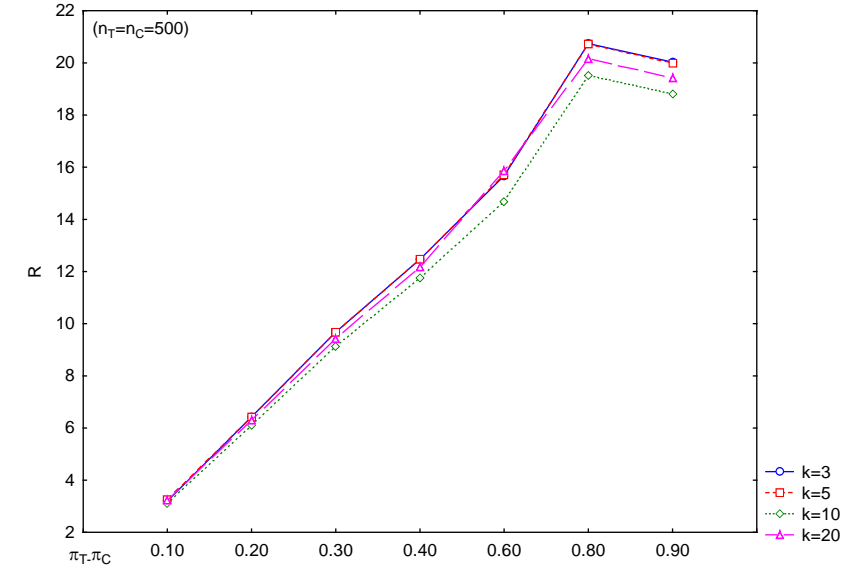
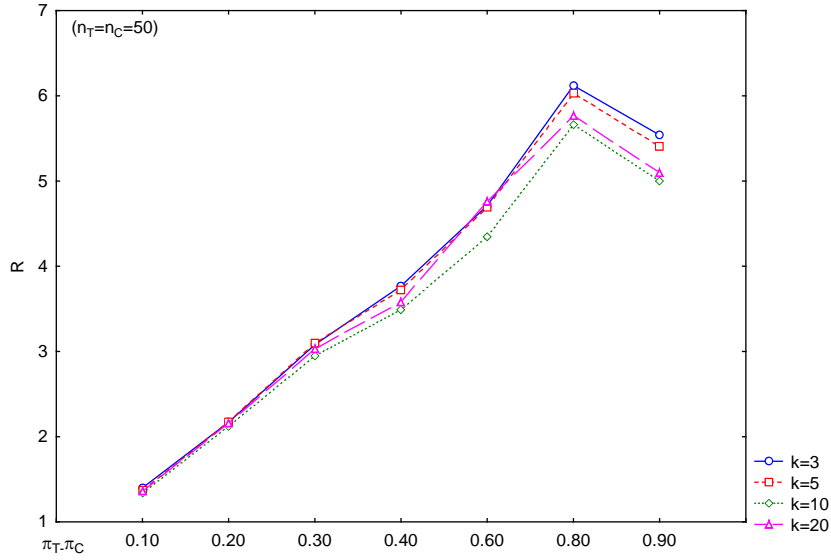
Şekil 4.9: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen H



Şekil 4.10: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen H_M



Şekil 4.11: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen τ^2



Şekil 4.12: Farklı örneklem büyüklüklerinde ve farklı tedavi etkinliklerinde elde edilen R

5.TARTIŞMA

Heterojenliğin test edilmesi, Cochran Q test istatistiğinin düşük güce sahip olmasından ya da heterojenlik ölçümü miktarının önemsiz bulunmasından dolayı yanlış anlaşılmaktadır. Q istatistiğinin gücü meta analizine dahil edilen çalışma sayısından etkilenmekte ve çalışma sayısı arttıkça Q değeri de artmaktadır. Meta analizine dahil edilen çalışma sayısı az ise ($k < 20$) testin gücü düşük, fazla ise testin gücü yüksek olmaktadır (48). Bununla birlikte eğer meta analizine pek çok çalışma dahil edildiğinde heterojenlik miktarı çok küçük de olsa bu değer test edildiğinde istatistiksel olarak anlamlı kabul edilebilir. Böyle bir durumda heterojenliğin olup olmamasına da klinik olarak önemli olup olmamasına da karar vermek zorlaşacaktır. Bu yüzden sadece istatistiksel test ile değerlendirmek değil aynı zamanda meta analizi üzerinden heterojenliğin etkisini de tanımlamamız tavsiye edilmektedir (31,43,46,49). Bu amaçla, Cochran Q testi ve heterojenlik ölçümlerinin meta analizine dahil edilen çalışma sayısından ve örneklem büyüklüğünden etkilenme durumları araştırılmıştır.

Ayrıca, literatürde Cochran Q istatistiğinin anlamlılık seviyesi üzerine pek çok tartışma yer almakta ve anlamlılık seviyesi için 0.05 yerine 0.10 alınması önerilmektedir (65). Bu duruma açıklık getirmek amacıyla, heterojenlik testine ait anlamlılık seviyesi ve heterojenlik ölçümleri için en uygun kesim noktalarının (cut-off) belirlenmesine çalışılmıştır.

Bu amaçla, yapılan simülasyon çalışmasında homojen ve heterojen yapıda oluşturulan kombinasyonlar göz önünde bulundurularak, 4 farklı çalışma sayısı (3, 5, 10 ve 20) ve 2 farklı örneklem büyüklüğü (50 ve 500) ele alınmış ve 56 homojen, 56 heterojen yapıda olmak üzere toplam 112 kombinasyon sonucu incelenmiştir. Bu sonuçlardan yola çıkarak, Cochran Q testi ve heterojenlik testlerinin çalışma sayısından ve örneklem büyüklüğünden etkilenme durumları grafiklerle ortaya konmaya çalışılmış, heterojenlik testine ait anlamlılık seviyesi için ve heterojenlik ölçümleri için ROC analizleri yapılmış, tanı testleri hesaplatılmış ve en uygun kesim noktaları elde edilmiştir.

5.1. Heterojenlik Testi ve Heterojenlik Ölçümlerinin Meta Analizine Dahil Edilen Çalışma Sayısından ve Örneklem Büyüklüğünden Etkilenme Durumları

Meta analizi çalışmalarında heterojenliğin belirlenmesinde çok yaygın olarak Cochran Q testi kullanılmaktadır. Cochran Q testi sonucuna göre, kullanılacak olan modele ve heterojenlik ölçümüne karar verilir. Cochran Q testi ve heterojenlik ölçümleri meta analizine

dahil edilen çalışma sayısından ve örneklem büyüklüğünden etkilenmektedir. Bu amaçla homojen, heterojen ve her ikisinin birlikte olduğu durumda oluşturulan kombinasyonlar dikkate alınarak, Cochran Q testi ve heterojenlik ölçümlerinin, çalışma sayısı ve örneklem büyüklüğü ile aralarındaki korelasyon katsayıları, korelasyon testine ait anlamlılık değerleri Çizelge 5.1’de, Cochran Q testi, heterojenlik ölçümleri ve % 95 güven aralıklarına ait grafik gösterimleri ise Şekil 5.1-5.6’da verilmektedir.

Çizelge 5.1: Cochran Q testi ve heterojenlik ölçümlerinin k ve N’ye göre korelasyon katsayıları ve p değerleri

		Cochran Q	I ²	H	H _M	R	τ ²
Homojen (n=56)	k	0.942 (<0.001)	-0.722 (<0.001)	-0.813 (<0.001)	-0.880 (<0.001)	-0.792 (<0.001)	-0.490 (<0.001)
	N	0.643 (<0.001)	-0.161 (0.235)	-0.303 (0.023)	-0.372 (0.005)	-0.265 (0.049)	-0.652 (<0.001)
Heterojen (n=56)	k	0.402 (0.002)	0.026 (0.849)	-0.018 (0.893)	-0.020 (0.881)	-0.019 (0.889)	-0.055 (0.686)
	N	0.606 (<0.001)	0.309 (0.020)	0.449 (0.001)	0.449 (0.001)	0.456 (<0.001)	-0.041 (0.767)
Homojen ve Heterojen birlikte (n=112)	k	0.268 (0.004)	-0.027 (0.774)	-0.013 (0.893)	-0.023 (0.809)	-0.013 (0.889)	-0.035 (0.716)
	N	0.401 (<0.001)	0.051 (0.593)	0.245 (0.009)	0.234 (0.013)	0.251 (0.008)	-0.026 (0.782)

Sonuçlar Cochran Q testine göre incelendiğinde, homojen durumda ve heterojen durumda benzer özellikler göstermekte olup, meta analizine dahil edilen çalışma sayısı ve örneklem büyüklüğü arttıkça doğrusal olarak Q istatistiği değerinin de arttığı görülmektedir (Şekil 5.1). Korelasyon tablosu incelendiğinde de, bu artışın her üç durumda da (homojen, heterojen ve tümü) olduğu ve istatistiksel olarak da anlamlı olduğu gözlenmektedir (Çizelge 5.1). Literatürde de heterojenlik testine ait benzer durumlar karşımıza çıkmaktadır. Cochran Q testinin bu probleminden dolayı pek çok heterojenlik ölçümleri ortaya atılmakta ve bu heterojenlik ölçümlerinin de heterojenlik testinin anlamlı olup olmamasına göre farklılık gösterdiği ifade edilmektedir. Özellikle çalışmalar arasında homojenliğin varlığında Cochran Q testinin yanı sıra I² ve H_M test istatistiğinin, heterojenliğin varlığında da çalışmalar arası varyans olarak adlandırılan τ² test istatistiğinin kullanımı önerilmektedir (31, 43, 45).

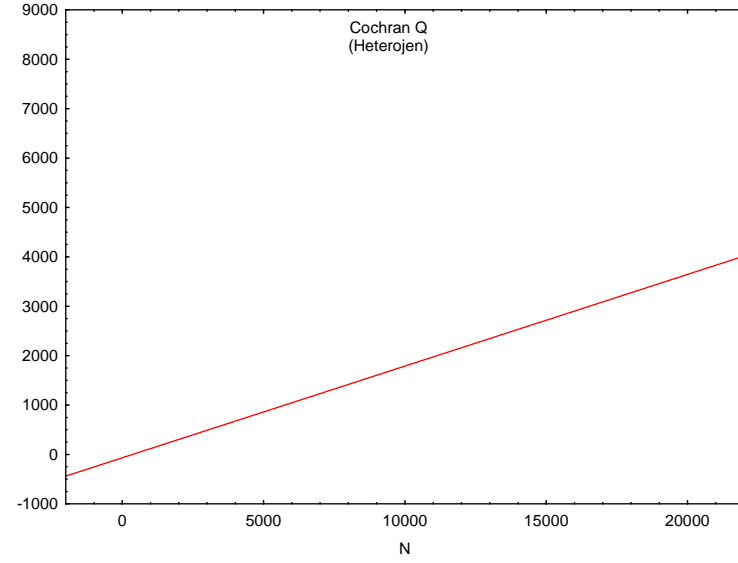
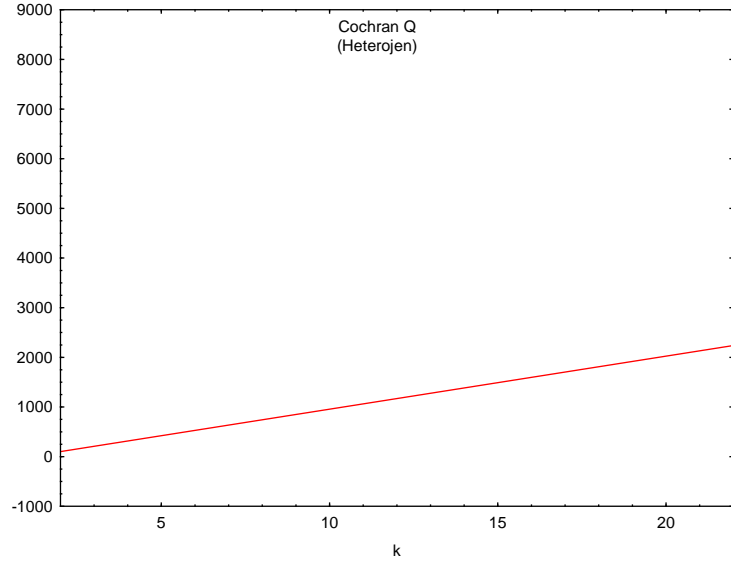
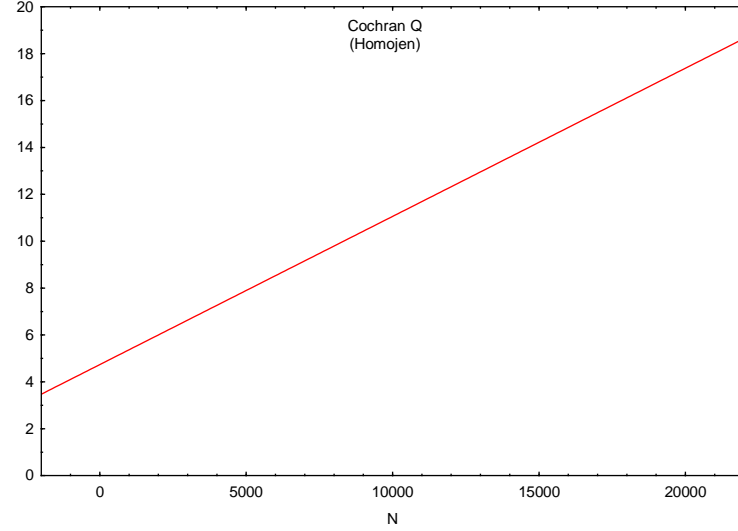
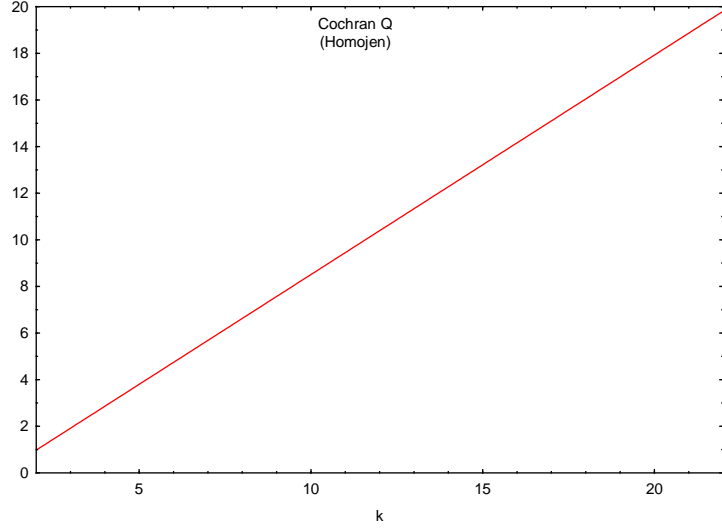
Sonuçlar I^2 test istatistiğine göre incelendiğinde, homojen durumda çalışma sayısı arttıkça I^2 istatistiğinin değerinin azaldığı, heterojen durumda çalışma sayısı arttıkça I^2 istatistiğinin değerinin de arttığı gözlenmektedir. Homojen ve heterojenin birlikte olduğu durumlar dikkate alındığında ise çalışma sayısından ve örneklem büyüklüğünden etkilenmemektedir (Çizelge 5.1). Şekil 5.2 incelendiğinde, homojen durumda çalışma sayısından etkilendiği çok net olarak görülmektedir. Çünkü I^2 istatistiğine ait üst sınır değeri çok büyük olduğundan dolayı çok geniş bir dağılım göstermektedir. Alt sınır değerine göre değişimi çok orantısız bir şekilde olmakta ve simetriklik bozulmaktadır. Medina ve ark. (2006), homojenlik kontrolü yapılırken, I^2 istatistiğine ait % 95 güven aralıklarının, meta analizine dahil edilen çalışma sayısının küçük olduğu durumlarda, Q testindeki benzer problemlerle karşılaşılacağını ifade etmektedir. Yani, Cochran Q testine ait eksiklikleri I^2 test istatistiğine ait güven aralıkları da çözememektedir. Bu nedenle, meta analizinde kullanılacak olan istatistiksel modele karar vermede Cochran Q testi ve I^2 indeksine ait güven aralıkları yanıltıcı olabilmektedir. Çalışma sayısı küçük olduğunda ($k < 20$), I^2 test istatistiğine ait güven aralığını ve Cochran Q testini yorumlarken çok dikkatli olunmalıdır. Güven aralığındaki bu probleme rağmen, meta analizine dahil edilen çalışma sayısının küçük olması durumunda heterojenliğin değerlendirilmesinde I^2 test istatistiğinin değerinin ve güven aralığının kullanılması önerilmektedir (31).

Sonuçlar H, H_M ve R test istatistiklerine göre incelendiğinde, H, H_M ve R test istatistiklerinin sonuçlarının benzer olduğu, homojen durumda çalışma sayısı ve örneklem büyüklüğü arttıkça H, H_M ve R test istatistiklerine ait değerlerin azaldığı, heterojen durumda da örneklem büyüklüğü arttıkça H, H_M ve R test istatistiklerine ait değerlerin arttığı gözlenmektedir. Homojen ve heterojen birlikte ele alındığında ise meta analizine dahil edilen çalışma sayısından değil örneklem büyüklüğünden etkilendikleri görülmektedir (Çizelge 5.1, Şekil 5.3-5.5). H, H_M ve R test istatistiklerine ait % 95 güven aralıkları incelendiğinde ise, I^2 test istatistiğinin güven aralığına ait sonuçlar ile benzer sonuçlar gösterdiği gözlenmektedir.

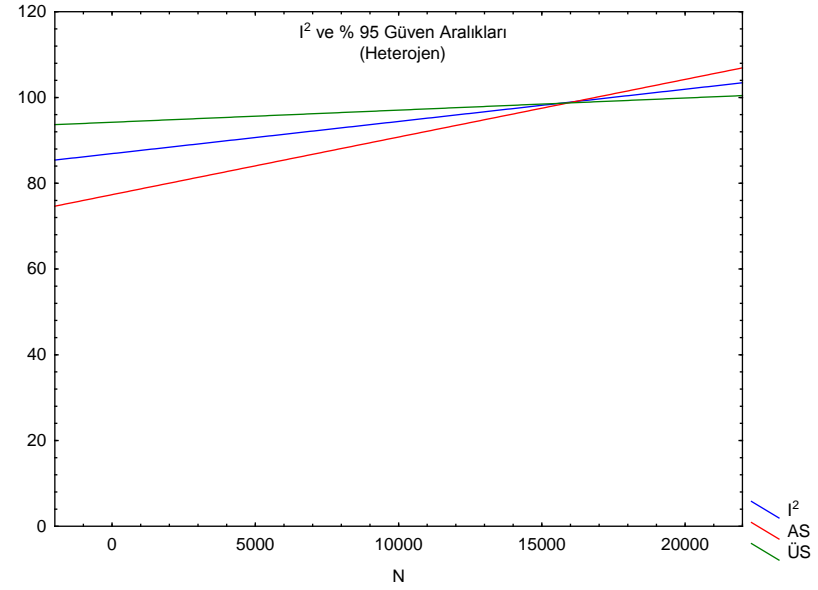
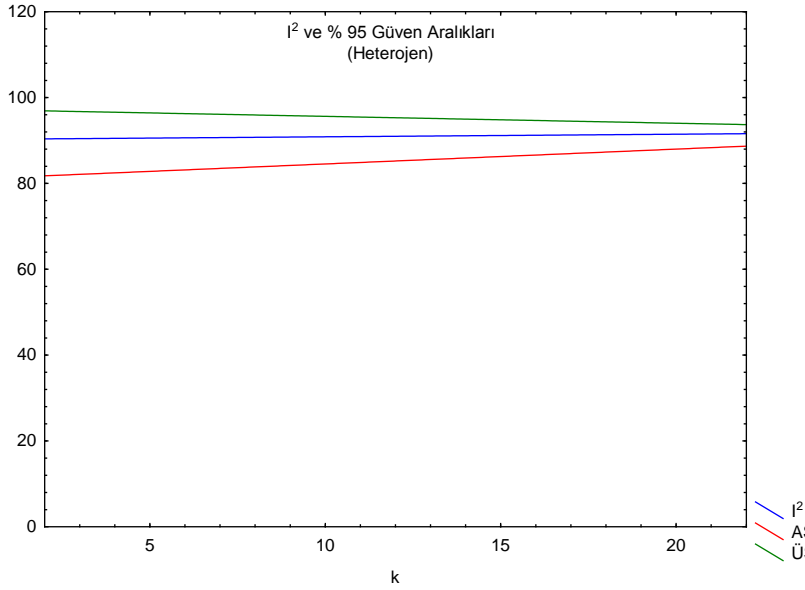
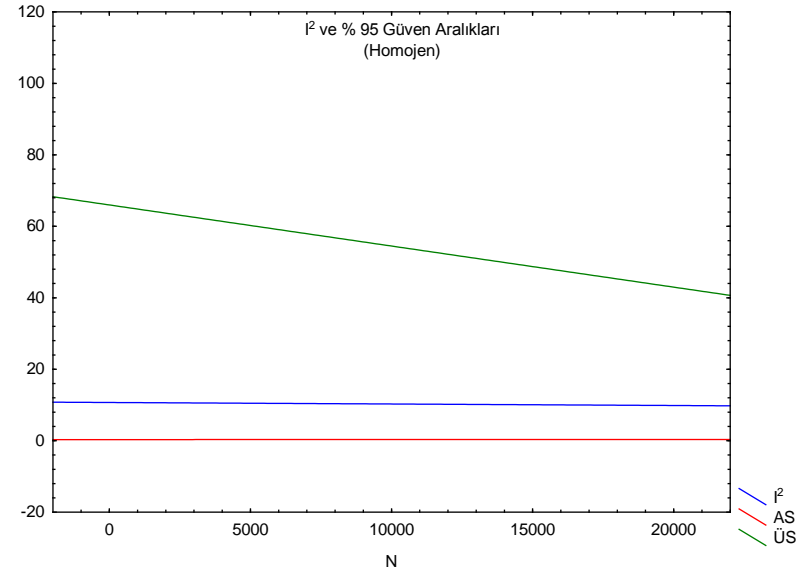
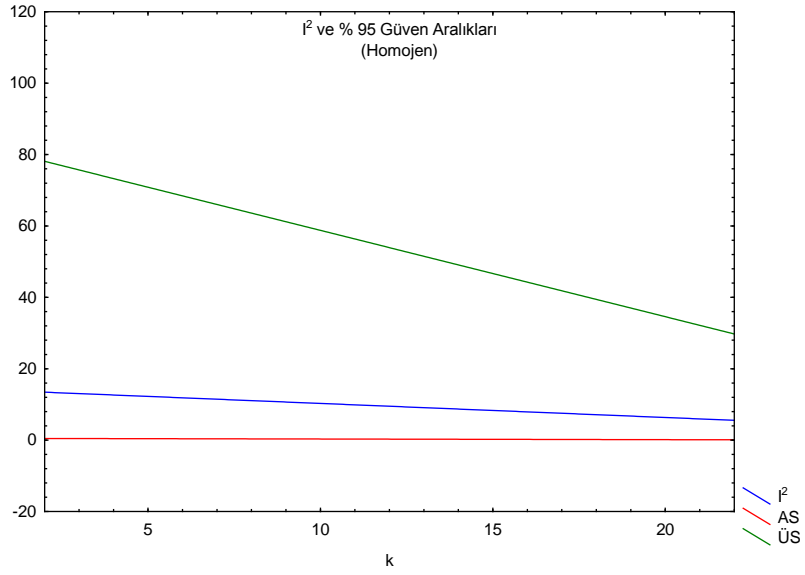
Sonuçlar τ^2 test istatistiğine göre incelendiğinde, homojen yapıda, çalışma sayısındaki ve örneklem büyüklüğündeki artışa bağlı olarak azaldığı, heterojen durumda ve tüm durumlarda çalışma sayısından ve örneklem büyüklüğünden etkilenmediği gözlenmektedir (Çizelge 5.1, Şekil 5.6). Güven aralığı bakımından incelendiğinde, heterojen durumda çalışma sayısındaki ve örneklem büyüklüğündeki artışa bağlı olarak daralan ve simetrik bir yapıya sahip olduğu gözlenmektedir (Şekil 5.6). Rucker ve ark. (2008), heterojenlik ölçümlerinden τ^2 test istatistiğinin, Cochran Q testi, I^2 , H, H_M ve R heterojenlik ölçümlerinin aksine örneklem büyüklüğünden etkilenmediğini, bu nedenle, örneklem büyüklüğü artırılırsa bu heterojenlik

ölçüm değerlerine ait güven aralıklarının daralacağını ve heterojenlik ölçümlerinin artacağını ifade etmektedir (46). Simülasyon bulgularına göre, sadece τ^2 test istatistiğinin değil aynı zamanda I^2 test istatistiğinin de çalışma sayısından ve örneklem büyüklüğünden etkilenmediğini söyleyebiliriz.

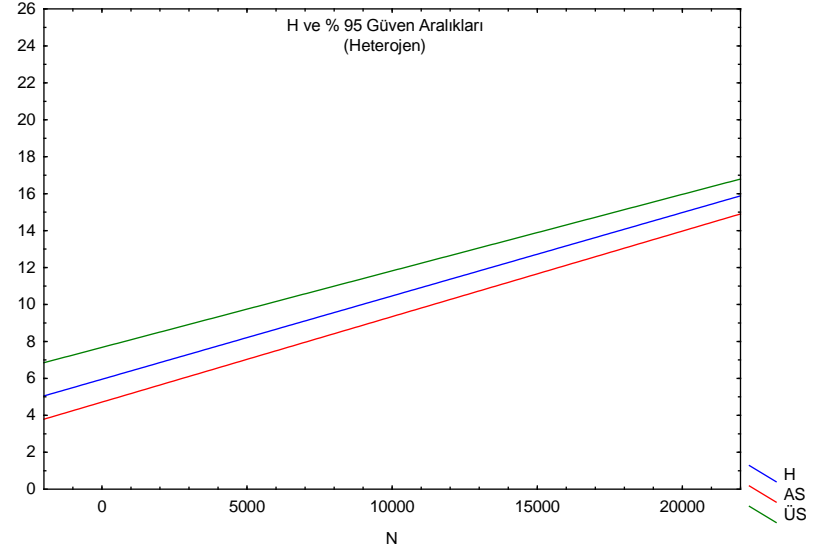
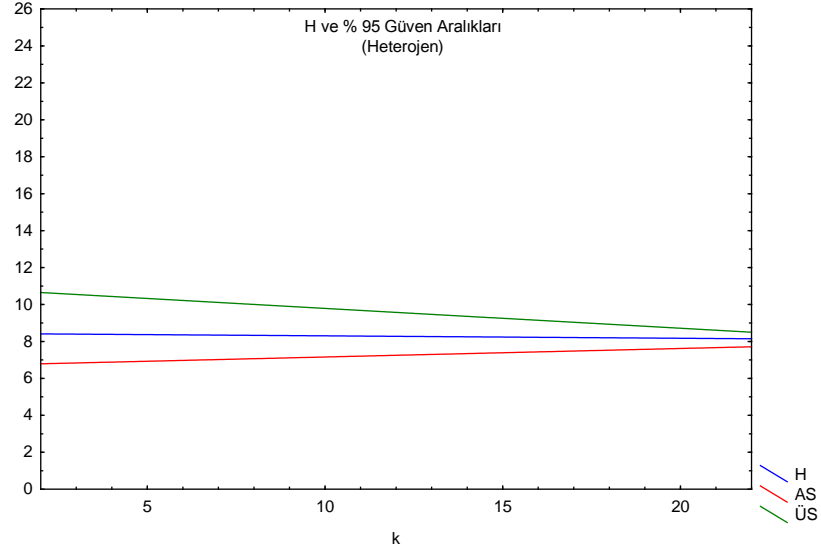
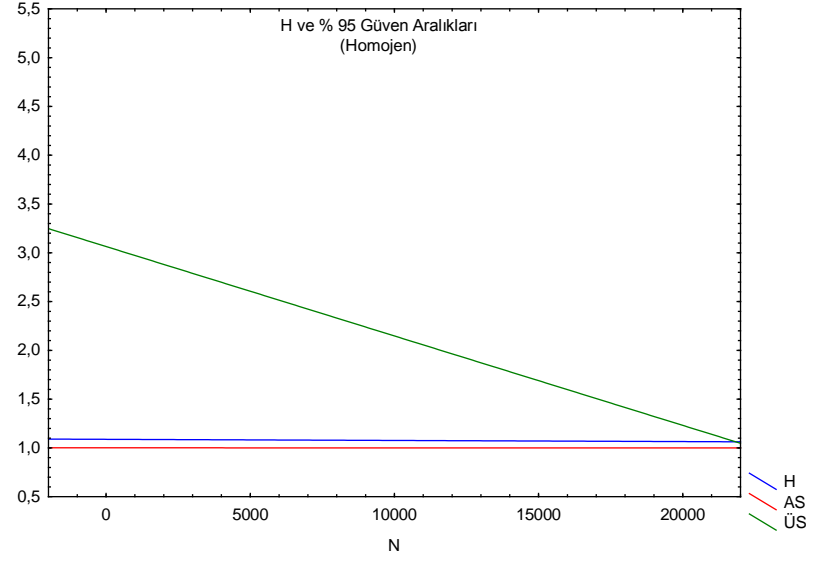
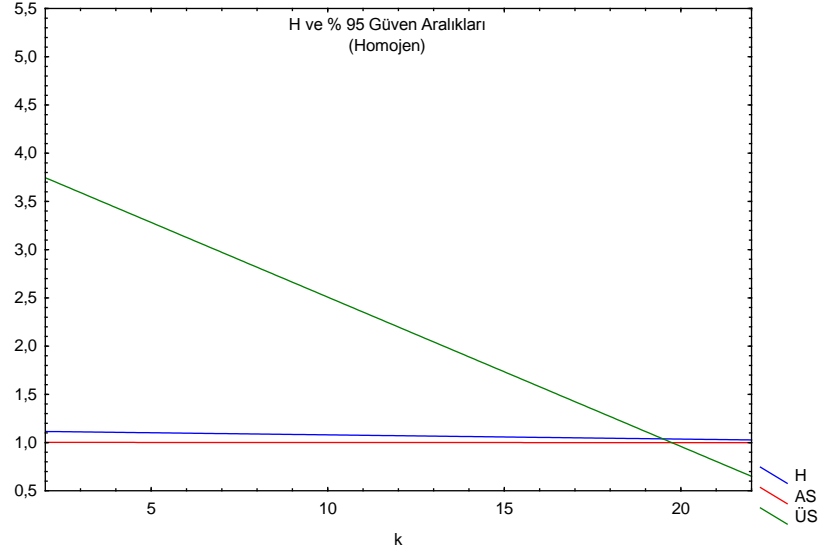
Meta analizi çalışmalarında heterojenlik ölçümlerinin, çalışma sayısından ve örneklem büyüklüğünden etkilenmemesi beklenmektedir. Bu simülasyon sonuçlarına göre, meta analizine alınacak olan çalışmaların, homojen veya heterojen olması, diğer bir deyişle, heterojenliğin test edilmesinde sabit etki veya rasgele etki modelinin tercih edilmesi durumlarında, heterojenlik ölçümü olarak I^2 ve τ^2 test istatistiğinin kullanılması önerilmektedir. Ancak bu heterojenlik ölçümlerinin güven aralıklarına ait problemler yaşanmaktadır.



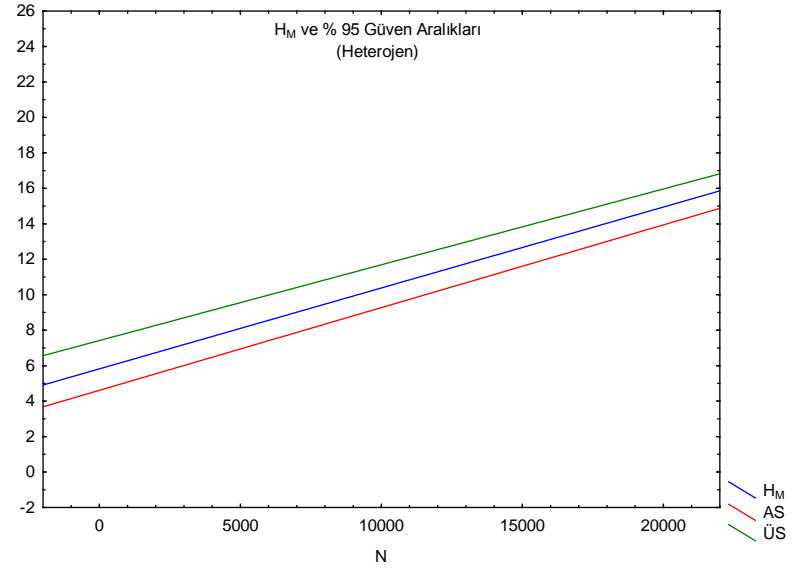
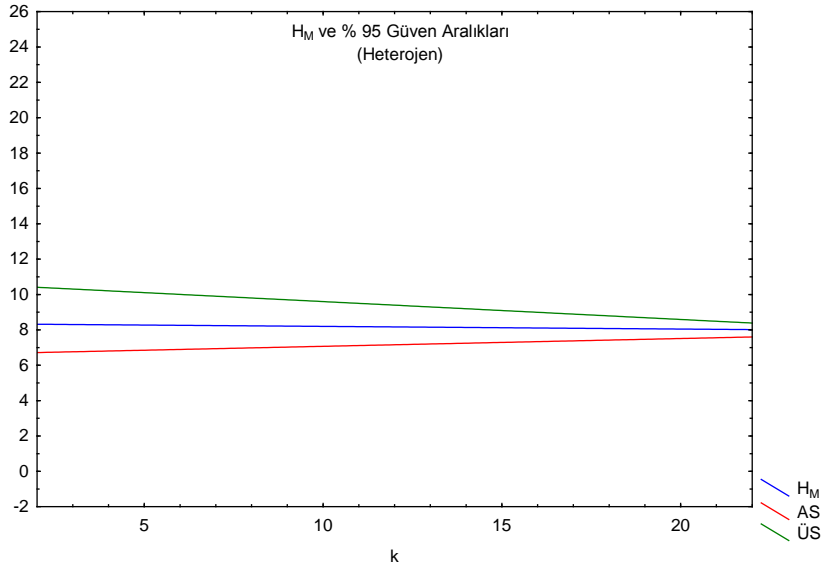
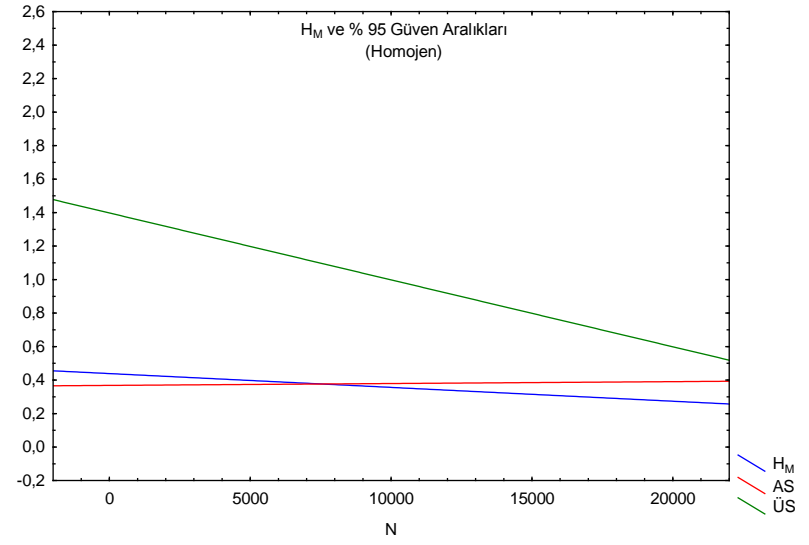
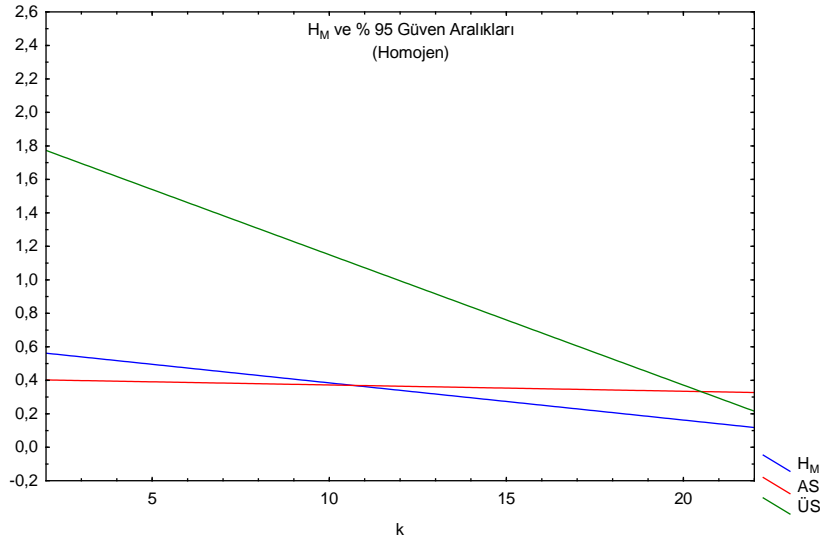
Şekil 5.1. Cochran Q'nun, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüğünden etkilenme durumu



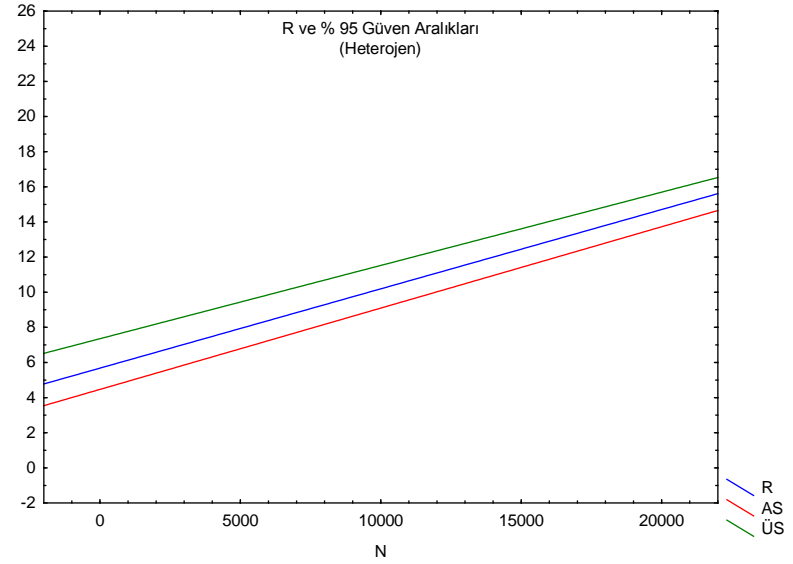
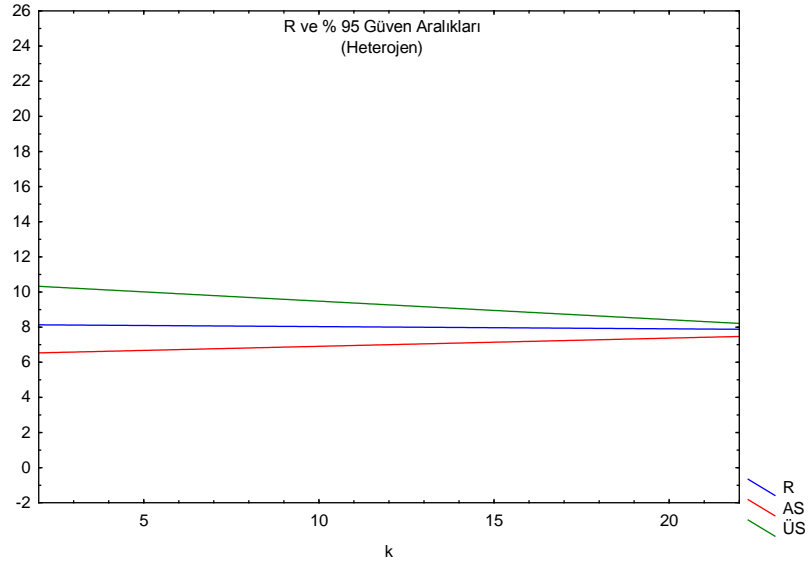
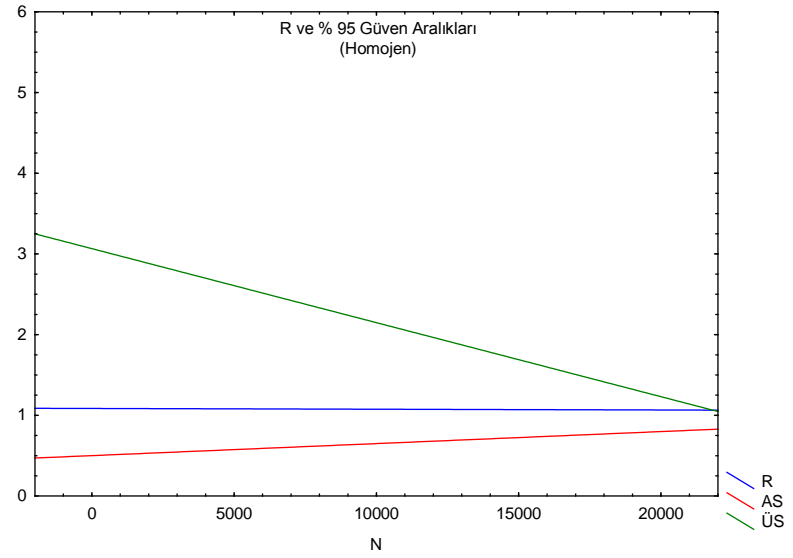
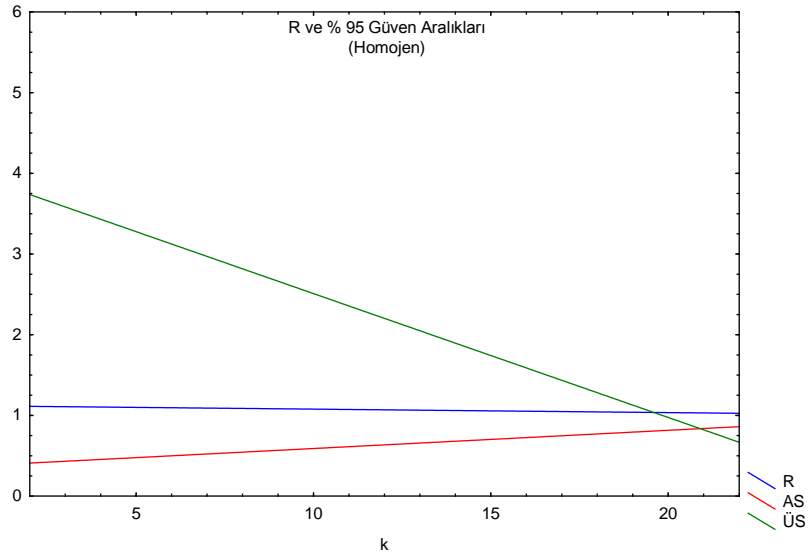
Şekil 5.2. I² test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu



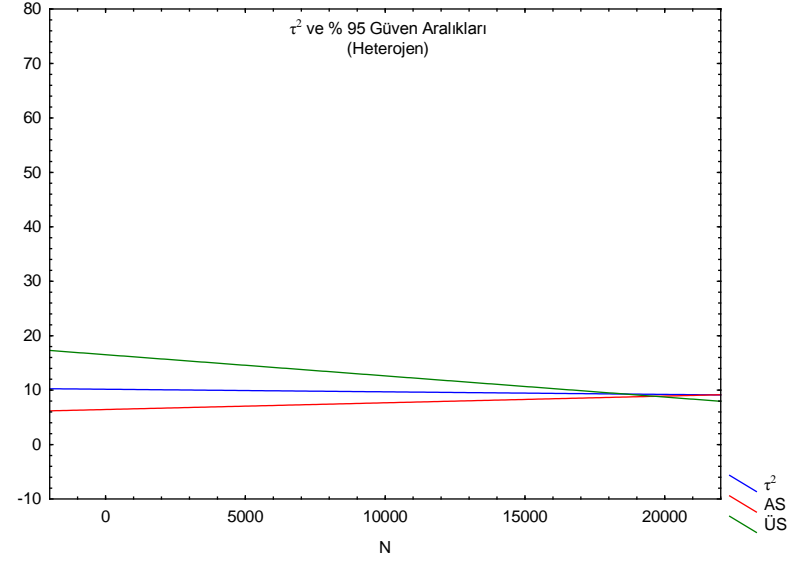
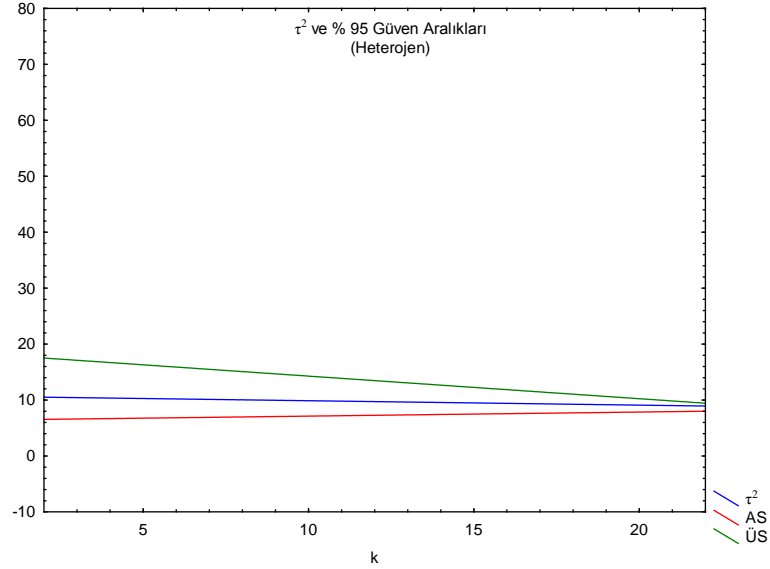
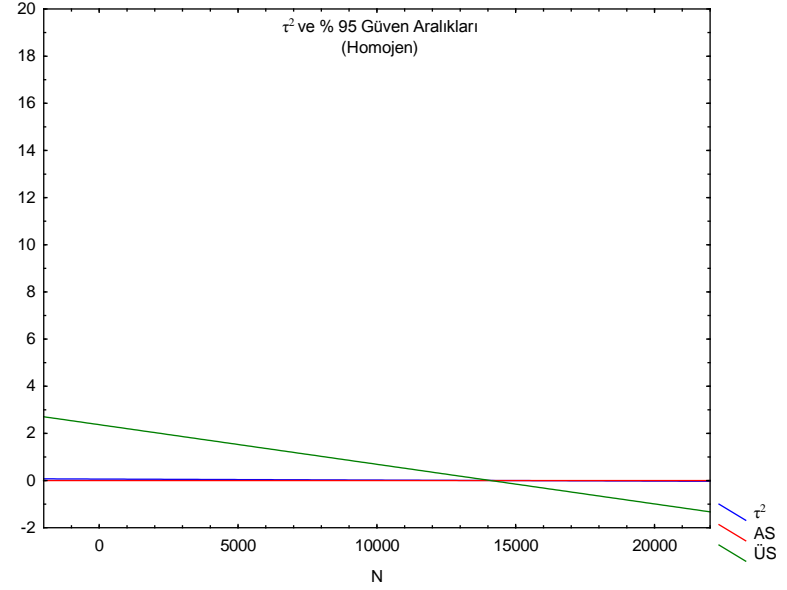
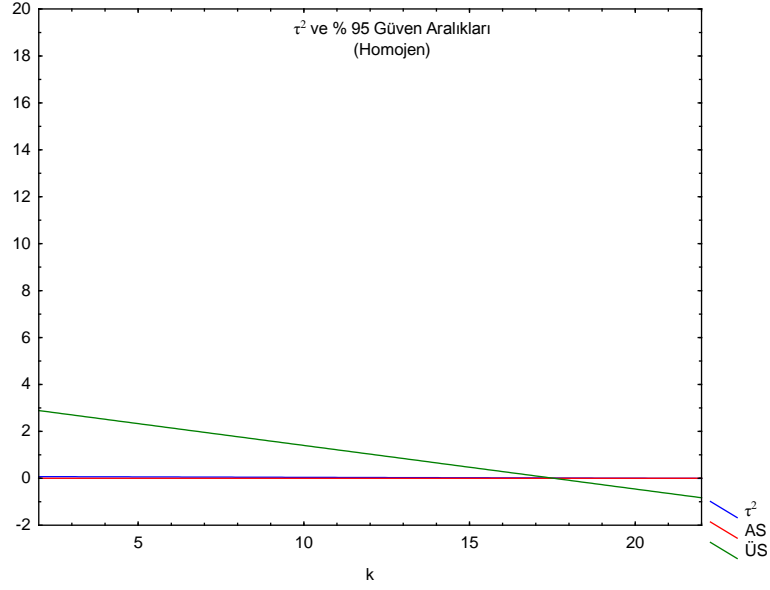
Şekil 5.3. H test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu



Şekil 5.4: H_M test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilene durumu



Şekil 5.5: R test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu



Şekil 5.6: τ^2 test istatistiğinin, homojen ve heterojen yapıda, çalışma sayısından ve örneklem büyüklüğünden etkilenme durumu

5.2. Heterojenlik Testine Ait Anlamlılık Seviyesi ile Heterojenlik Ölçümlerine Ait Kesim Noktalarının Belirlenmesi

Meta analizi çalışmalarında en uygun heterojenlik ölçümü, çalışma sayısından ve örneklem büyüklüğünden etkilenmemelidir. Simülasyon bulgularına göre, çalışma sayısından ve örneklem büyüklüğünden etkilenmeyen heterojenlik ölçümleri I^2 ve τ^2 istatistikleridir. Ancak, bu heterojenlik ölçümlerinin güven aralıkları ile ilgili problemler olduğu gözlenmektedir. Buradan yola çıkarak, meta analizi çalışmalarında kullanılacak olan heterojenlik testinin ve heterojenlik ölçümlerinin I. ve II. Tip hatadan etkilenme durumlarının ortaya konması ve heterojenlik ölçümlerine ait kesim noktalarının belirlenmesinin gerekliliği sonucuna varılmıştır. Literatürde Cochran Q istatistiğinin anlamlılık seviyesi üzerine pek çok tartışma yer almakta ve anlamlılık seviyesi için 0.05 yerine 0.10 alınması önerilmektedir. Bu amaçla, homojen ve heterojen yapıda oluşturulan kombinasyonlar göz önünde bulundurularak heterojenlik testine ait anlamlılık seviyesi ve heterojenlik ölçümleri için ROC analizleri yapılmış ve uygun kesim noktaları elde edilmiştir. Ayrıca, simülasyon bulgularına göre, literatürde Cochran Q heterojenlik testinin, 0.05 ve 0.10 anlamlılık seviyeleri için de heterojenlik ölçümlerine ait kesim noktaları elde edilmiştir. 0.05, 0.10 ve 0.30 anlamlılık seviyelerinde, çalışma sayısına bağlı olarak heterojenlik ölçümlerinin gerçek homojen ve heterojen durumlardaki değişimleri grafiklerle verilmiştir.

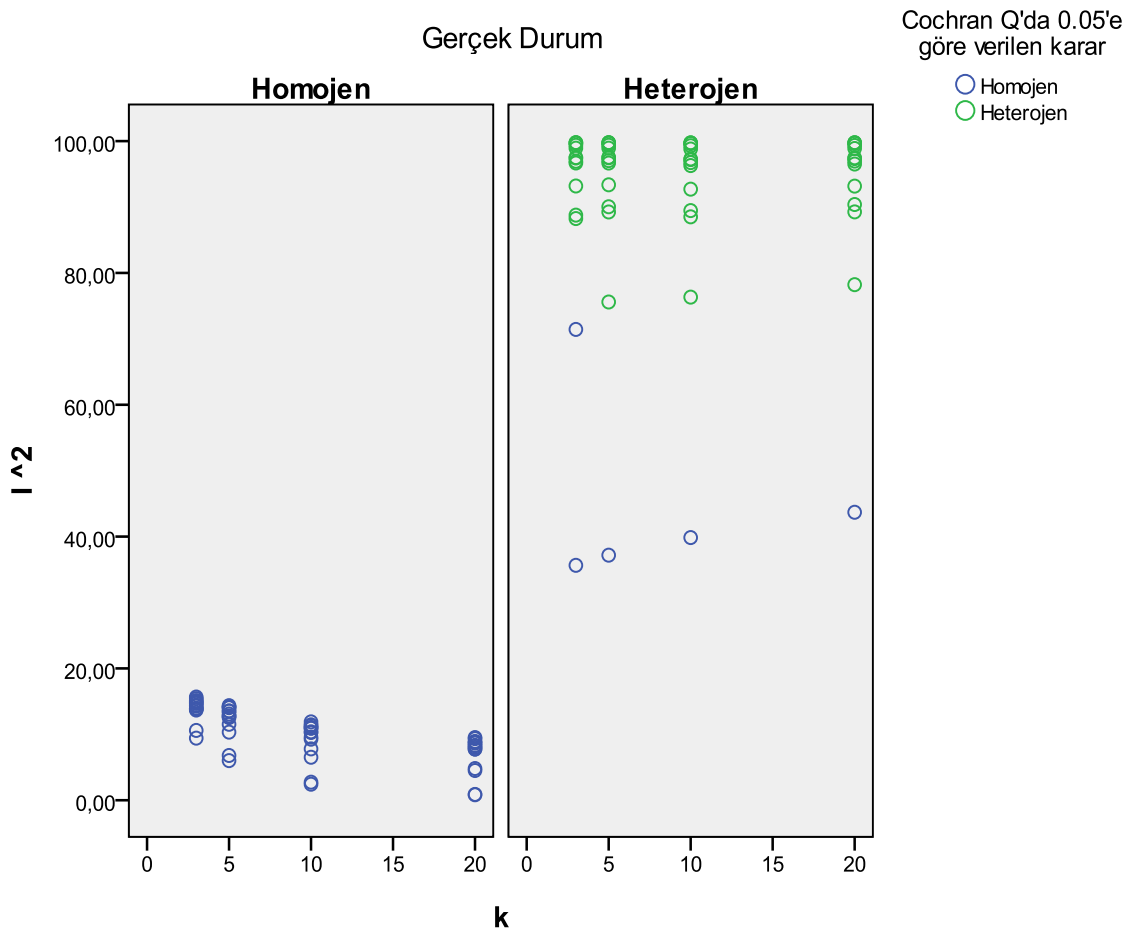
Cochran Q testine göre anlamlılık seviyesi 0.05 olarak ele alındığında, heterojenlik ölçümlerine ait kesim noktaları Çizelge 5.2'de, anlamlılık seviyesi 0.10 olarak alındığında heterojenlik ölçümlerine ait kesim noktaları Çizelge 5.3'de, homojen ve heterojen yapıda oluşturulan kombinasyonlardan elde edilen durumda, heterojenlik testine ait anlamlılık seviyesi ve heterojenlik ölçümlerinin kesim noktaları ise Çizelge 5.4'de verilmektedir.

Cochran Q testine göre anlamlılık seviyesi 0.05 olarak ele alındığında, heterojenlik ölçümü olan I^2 test istatistiği için kesim noktası % 71.43, H için 1.41, H_M için 1.1, R için 1.4 ve τ^2 için ise 0.24 olarak elde edilmekte ve bu değerlerin üzerindeki değerler için meta analizine dahil edilen çalışmaların heterojen olduğu sonucuna varılmaktadır (Çizelge 5.2). Meta analizine dahil edilen çalışmaların gerçekte heterojen olması gerekirken, 0.05 anlamlılık seviyesine göre karar verildiğinde I^2 , H, H_M ve R test istatistiklerinin bazı değerleri homojen olarak yorumlanmaktadır (Şekil 5.7-5.10). Yalnızca çalışmalar arası varyans olarak da bilinen τ^2 heterojenlik ölçüm değerleri, Cochran Q testinde 0.05 anlamlılık seviyesine göre karar verilirken gerçek homojen ve heterojenleri çok iyi ayırt edebilmektedir (Şekil 5.11). Bu sonuçlara göre denilebilir ki, Cochran Q test istatistiğinde 0.05 anlamlılık seviyesine göre

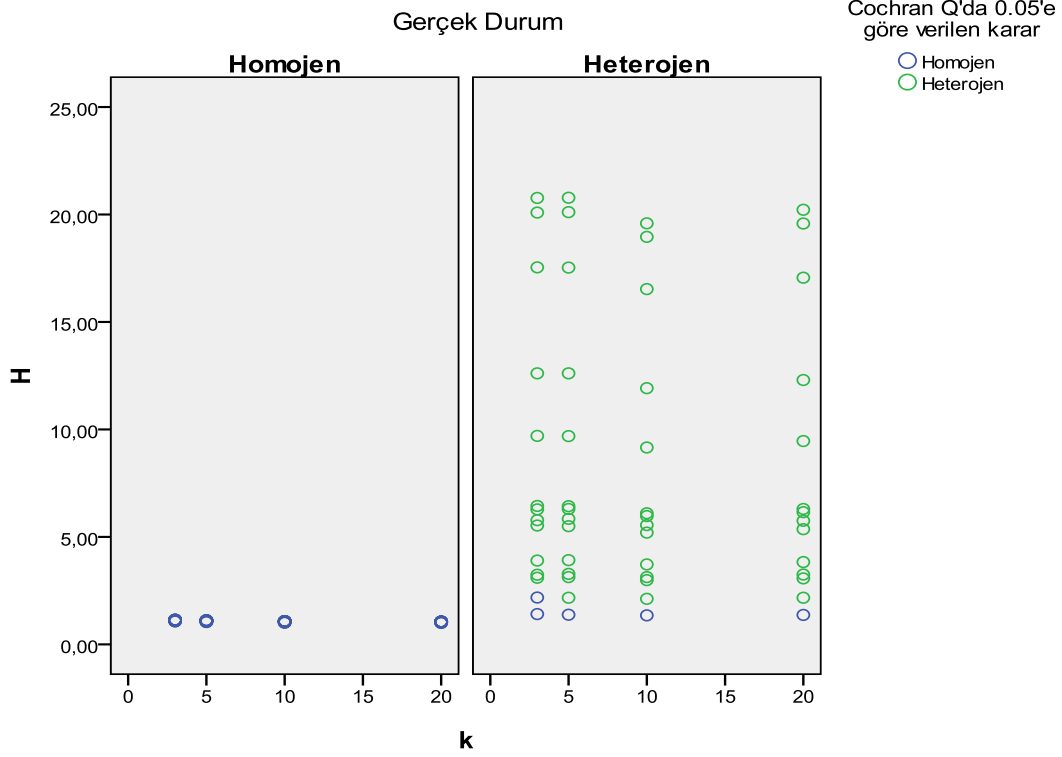
karar verilmek isteniyorsa heterojenlik ölçümü olarak I^2 heterojenlik ölçümü kullanılması önerilebilir.

Çizelge 5.2: Cochran Q'ya göre 0.05 anlamlılık seviyesinde heterojenlik ölçümlerine ait kesim noktaları

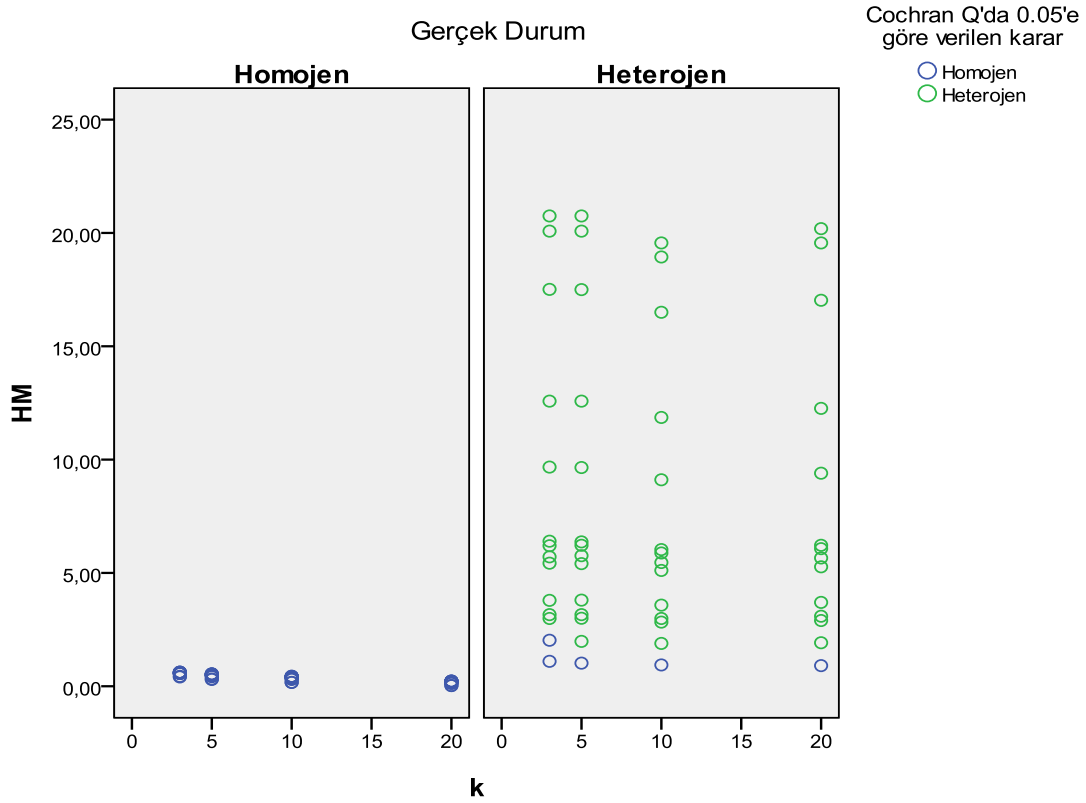
P<0.05	Duyarlılık	Seçicilik	LR(+)	LR(-)	AUC (p)	Kesim Noktası
I^2	100.00 (93.0-100.0)	100.00 (94.1-100.0)	0.00	0.00	1.000 (< 0.001)	> 71.43
H	100.00 (93.0-100.0)	98.36 (91.2-100.0)	61.00 (59.1-63.0)	0.00	0.999 (< 0.001)	> 1.41
H_M	100.00 (93.0-100.0)	98.36 (91.2-100.0)	61.00 (59.1-63.0)	0.00	0.999 (< 0.001)	> 1.1
τ^2	92.16 (81.1-97.8)	98.36 (91.2-100.0)	56.22 (51.6-61.3)	0.080 (0.009-0.7)	0.989 (< 0.001)	> 0.24
R	100.00 (93.0-100.0)	98.36 (91.2-100.0)	61.00 (59.1-63.0)	0.00	0.999 (< 0.001)	> 1.4



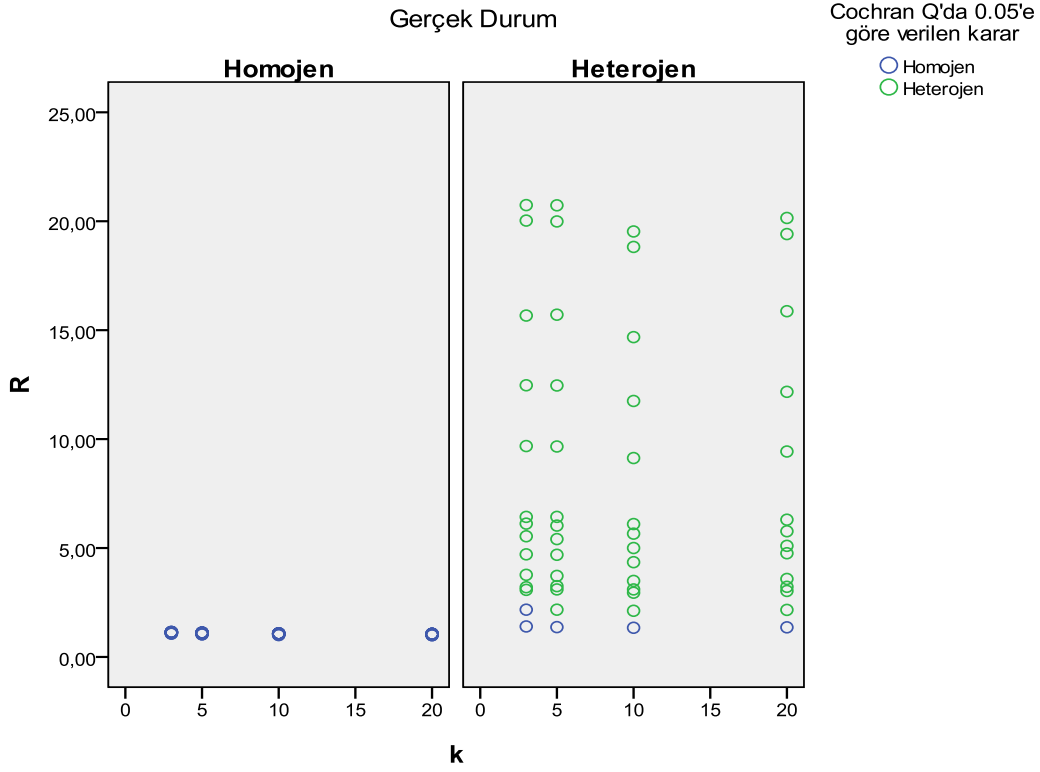
Şekil 5.7. I^2 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



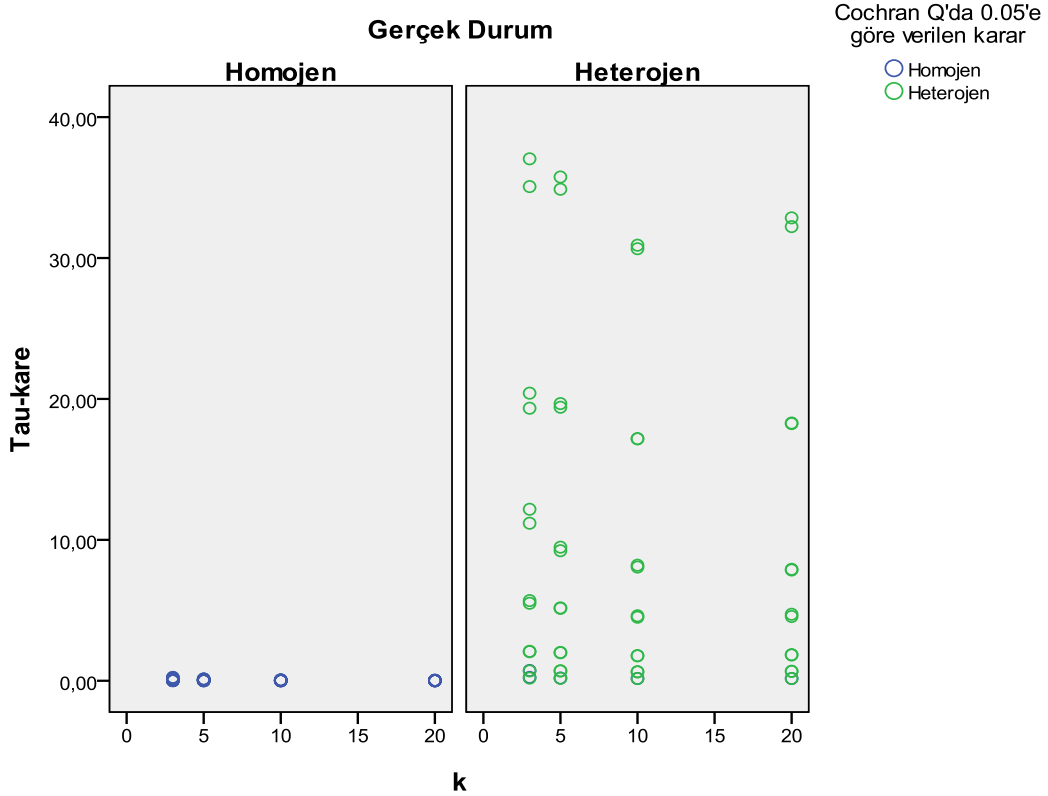
Şekil 5.8. H 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



Şekil 5.9. H_M 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



Şekil 5.10. R 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



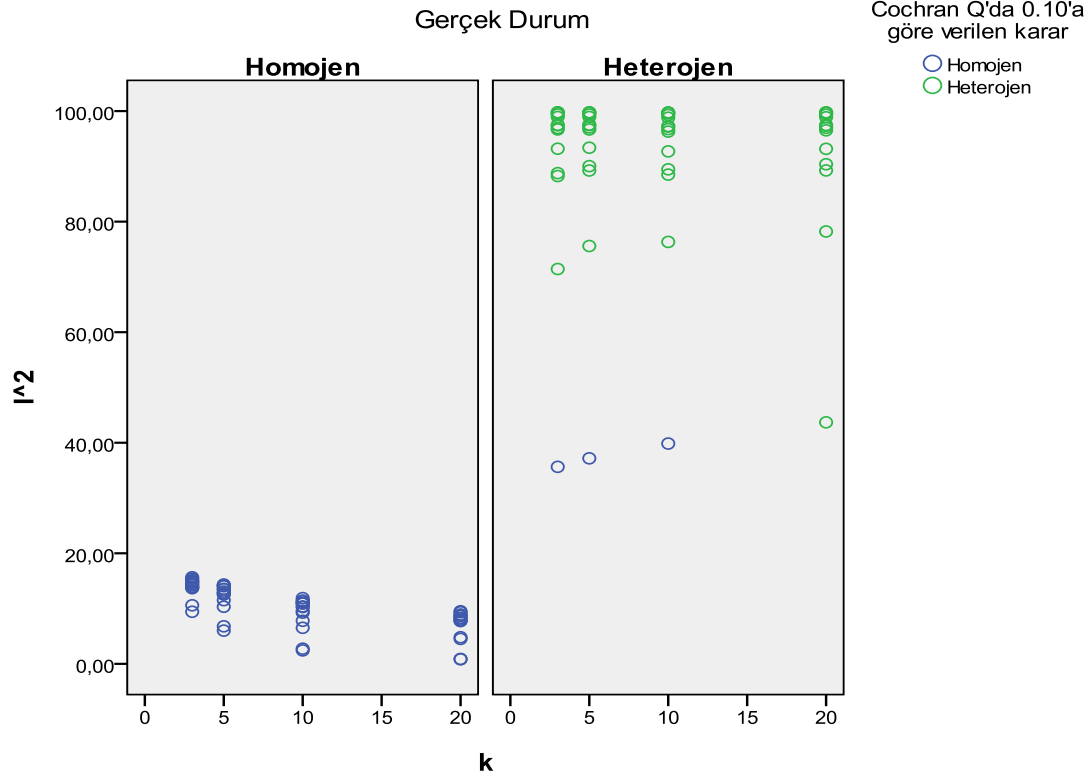
Şekil 5.11. τ^2 'nin, Cochran Q'da 0.05 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması

Cochran Q testine göre anlamlılık seviyesi 0.10 olarak ele alındığında heterojenlik ölçümlerine ait kesim noktaları ve literatürde yer alan kesim noktaları Çizelge 5.3'de verilmektedir. Heterojenlik ölçümlerinden I^2 değeri için literatürde % 50'nin üzerinde olduğunda çalışmaların heterojen olduğu sonucuna varılırken yapılan simülasyon sonucunda homojenlik sınırı % 39.85 olarak belirlenmiş ve bu değer üzerindeki değerler heterojen olarak kabul edilmiştir. Diğer heterojenlik ölçümü olan H ve R test istatistikleri için, 1 olduğunda mükemmel bir homojenlik sonucuna varılırken simülasyon sonucunda 1.4 olarak belirlenmiştir. Bu değer üzerindeki tüm değerler için çalışmalar heterojen olarak kabul edilirken, yanlılığı göz önünde bulunduran H_M heterojenlik ölçümüne göre 1.1'in üzerindeki değerler heterojen olarak kabul edilmektedir. Çalışmalar arası varyans olarak bilinen τ^2 için ise 0.24'ün üzerinde olduğu durumlarda meta analizine dahil edilen çalışmaların heterojen olduğu sonucuna varılmaktadır.

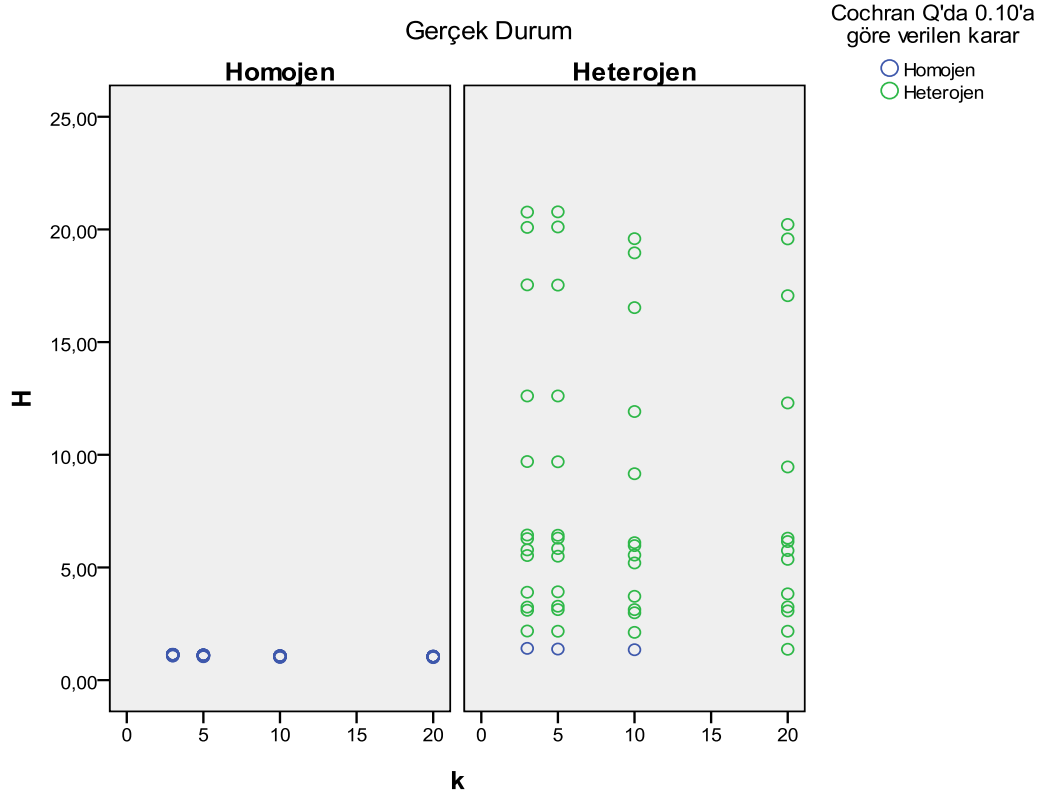
Meta analizine dahil edilen çalışmaların gerçekte heterojen olması gerekirken, 0.10 anlamlılık seviyesine göre karar verildiğinde I^2 , H, H_M ve R test istatistiklerinin bazı değerleri homojen olarak yorumlanmaktadır (Şekil 5.12-5.15). τ^2 heterojenlik ölçüm değerleri incelendiğinde, Cochran Q testinde 0.10 anlamlılık seviyesine göre karar verilirken gerçek homojen ve heterojenleri çok iyi ayırt edebilmektedir (Şekil 5.16). Bu sonuçlara göre denilebilir ki, Cochran Q test istatistiğinde 0.10 anlamlılık seviyesine göre karar verilmek isteniyorsa heterojenlik ölçümü olarak I^2 heterojenlik ölçümü kullanılması önerilebilir. Ayrıca, I^2 heterojenlik ölçümü hariç diğer heterojenlik ölçümlerinin kesim noktalarının anlamlılık seviyesi 0.05 ve 0.10'da değişmediği görülmektedir. Yani, Cochran Q heterojenlik testine göre karar verilirken anlamlılık seviyesi 0.05'de 0.10'da alınsa kesim noktalarının aynı olduğu gözlenmektedir.

Çizelge 5.3 Cochran Q'ya göre 0.10 anlamlılık seviyesinde heterojenlik ölçümlerine ait kesim noktaları

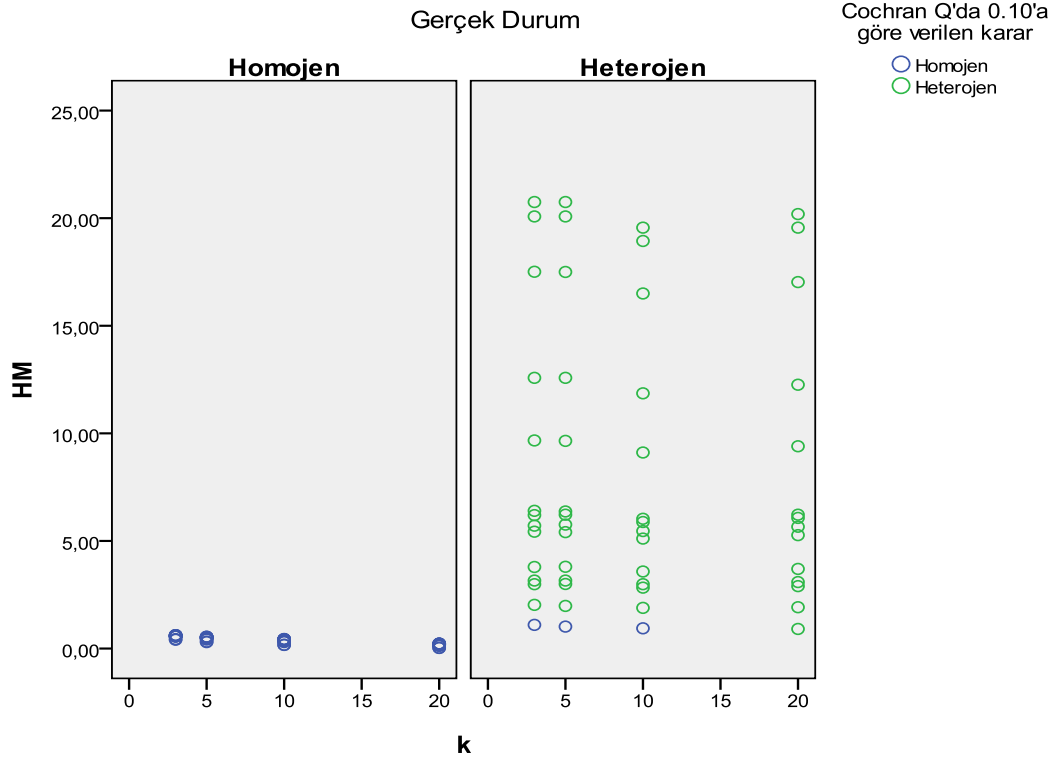
P<0.10	Duyarlılık	Seçicilik	LR (+)	LR(-)	AUC (p)	Kesim Noktası	Literatür Kesim Noktası (31,43,46,50)
I^2	100.00 (93.3-100.0)	100.00 (93.9-100.0)		0.00	1.000 (< 0.001)	> 39.85	> 50.0
H	98.11 (89.9-100.0)	100.00 (93.9-100.0)		0.019 (0.001-0.3)	0.999 (< 0.001)	> 1.41	> 1.00
H_M	98.11 (89.9-100.0)	100.00 (93.9-100.0)		0.019 (0.001-0.3)	0.999 (< 0.001)	> 1.1	> 0.0
τ^2	90.57 (79.3-96.9)	100.00 (93.9-100.0)		0.094 (0.02-0.5)	0.991 (< 0.001)	> 0.24	> 0.0
R	98.11 (89.9-100.0)	100.00 (93.9-100.0)		0.019 (0.001-0.3)	0.999 (< 0.001)	> 1.4	> 1.00



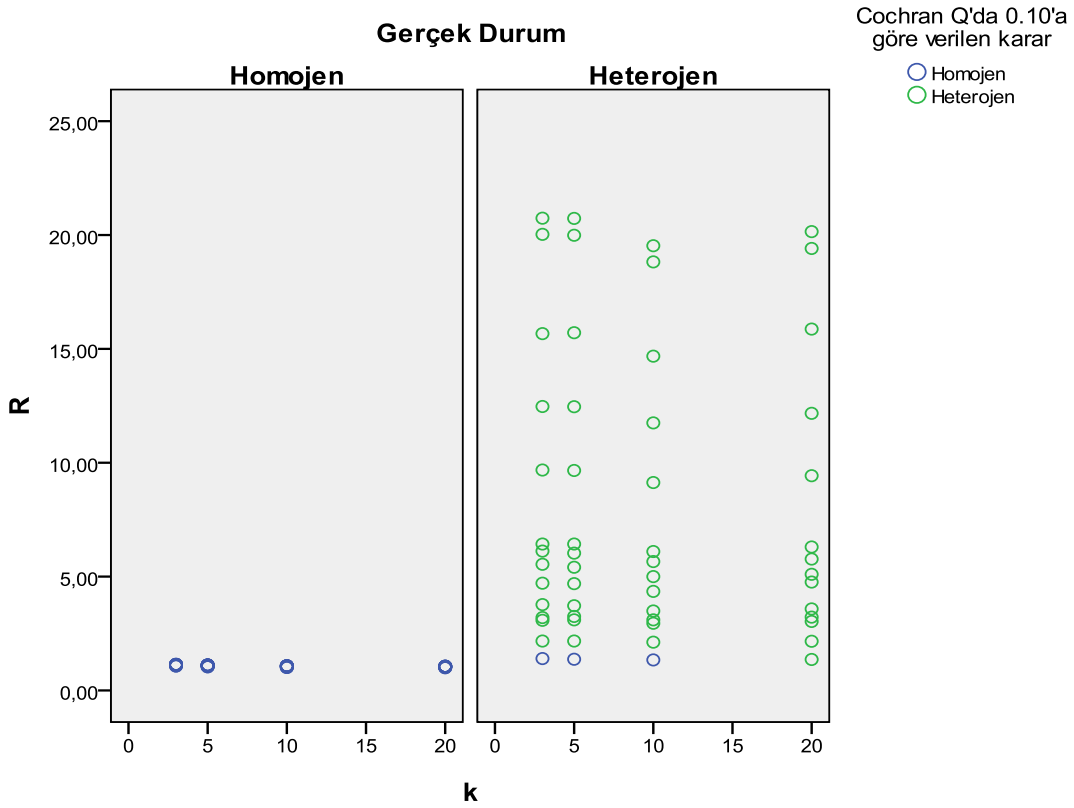
Şekil 5.12. I^2 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



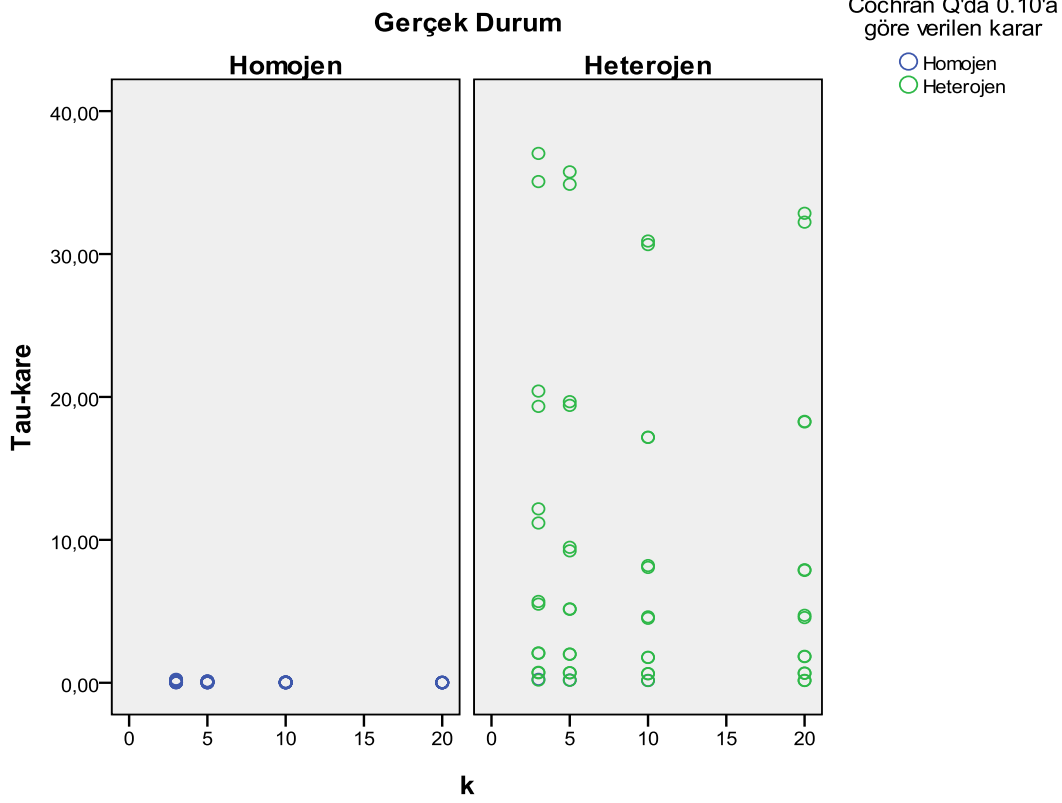
Şekil 5.13. H 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



Şekil 5.14. H_M 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



Şekil 5.15. R 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



Şekil 5.16. τ^2 'nin, Cochran Q'da 0.10 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması

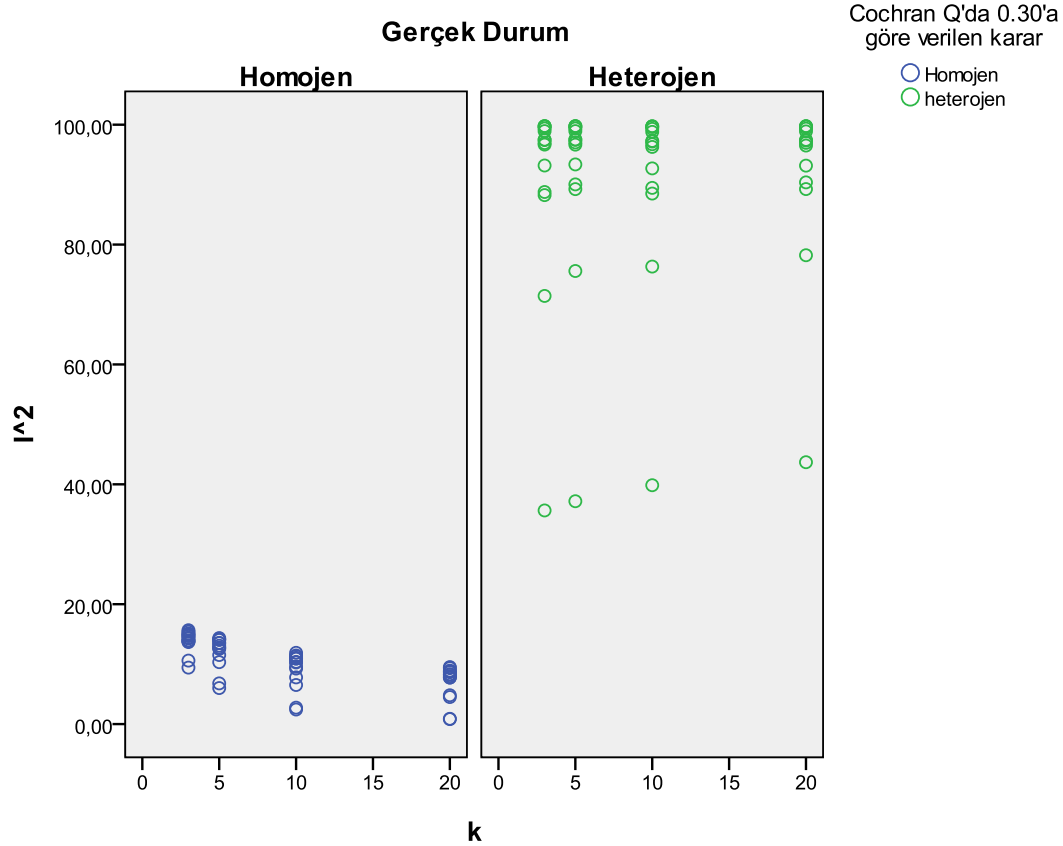
Homojen ve heterojen yapıda oluşturulan kombinasyonlar dikkate alındığında, Cochran Q heterojenlik testine ait anlamlılık seviyesi ve heterojenlik ölçümleri için kesim noktaları belirlenmiş ve bunlara ait sonuçlar Çizelge 5.4'de verilmiştir. Buna göre, meta analizine dahil edilen çalışmaların homojen olduğunu söyleyebilmek için heterojenlik testi sonucunda elde edilen anlamlılık seviyesi 0.30'un üzerinde olması gerekmektedir. Heterojenlik ölçümlerinden I^2 değeri % 50'nin üzerinde olduğunda çalışmaların heterojen olduğu sonucuna varılırken yapılan simülasyon sonucunda homojenlik sınırı % 15.66 olarak belirlenmiş ve bu değer üzerindeki değerler heterojen olarak kabul edilmiştir. Diğer heterojenlik ölçümü olan H ve R test istatistikleri için, 1.15'in üzerindeki tüm değerler için çalışmalar heterojen olarak kabul edilirken, yanlışlığı göz önünde bulunduran H_M heterojenlik ölçümüne göre 0.62'nin üzerindeki değerler heterojen olarak kabul edilmektedir. Çalışmalar arası varyans olarak bilinen τ^2 için ise 0.13'ün üzerinde olduğu durumlarda meta analizine dahil edilen çalışmaların heterojen olduğu sonucuna varılmaktadır.

Cochran Q testinde 0.30 anlamlılık seviyesine göre karar verilirken, tüm heterojenlik ölçümleri gerçek homojen ve heterojenleri çok iyi ayırt edebilmektedir (Şekil 5.17-5.21). Bu sonuçlardan yola çıkarak, meta analizi çalışmalarında, heterojenlik testinden yola çıkılarak

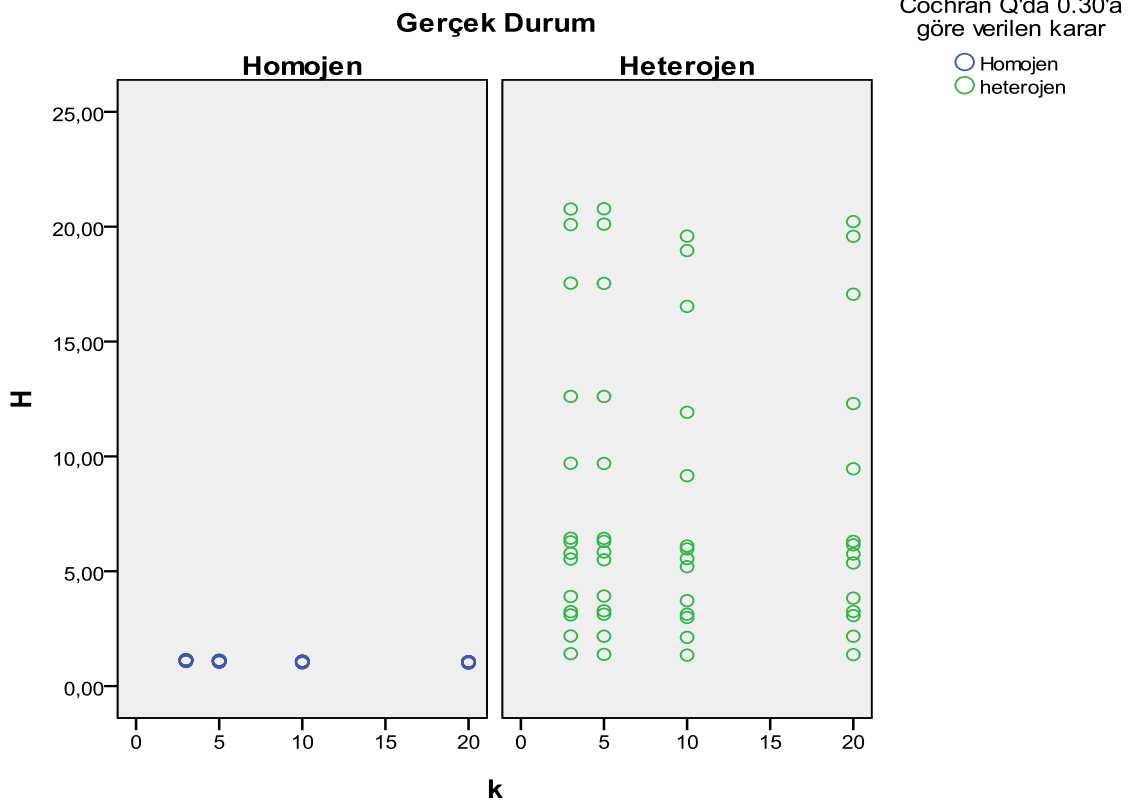
verilecek olan kararın doğru verilebilmesi için, Cochran Q testinde anlamlılık seviyesi olarak 0.30'un alınması önerilmektedir.

Çizelge 5.4: Gerçek durumda heterojenlik testine ait anlamlılık seviyesi ve heterojenlik ölçümleri için kesim noktaları

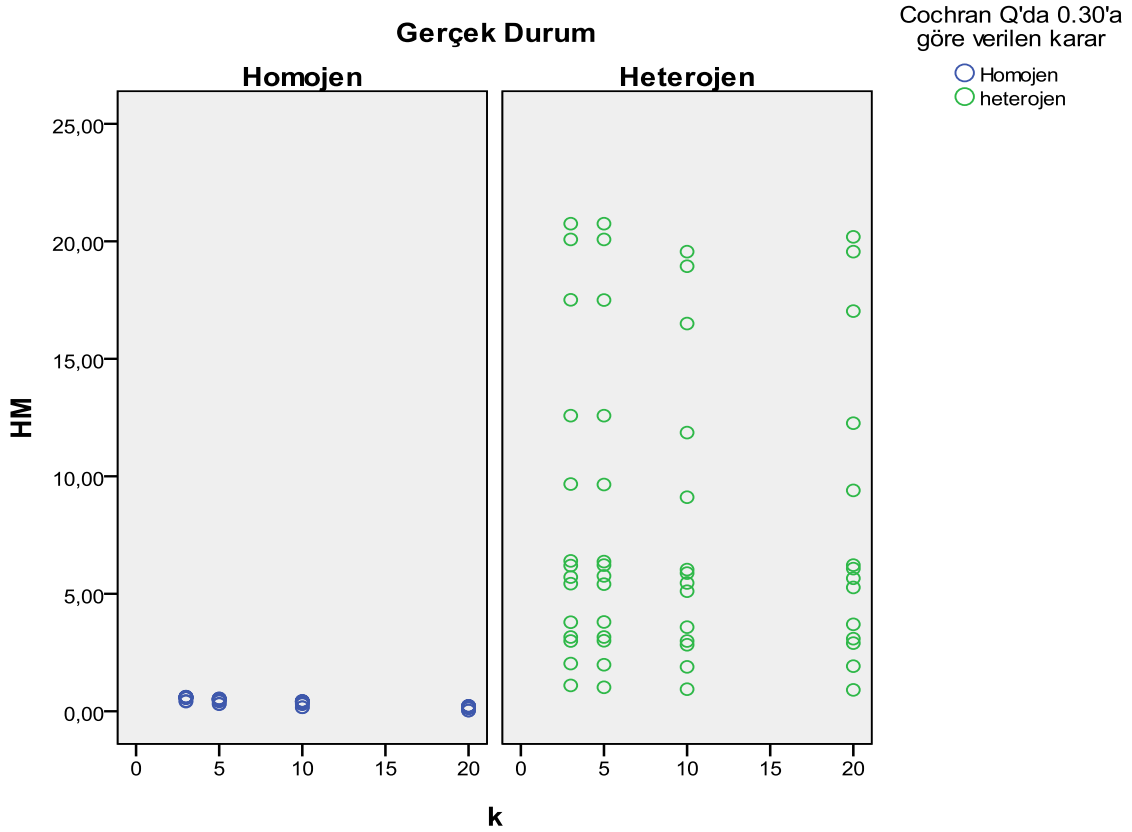
Gold standart	Duyarlılık	Seçicilik	LR(+)	LR(-)	AUC (p)	Kesim Noktası
P değeri	100.00 (93.6-100.0)	100.00 (93.6-100.0)		0.00	1.000 (< 0.001)	≤ 0.29
I^2	100.00 (93.6-100.0)	100.0 (93.6-100.0)		0.00	1.000 (< 0.001)	> 15.66
H	100.00 (93.6-100.0)	100.0 (93.6-100.0)		0.00	1.000 (< 0.001)	> 1.15
H_M	100.00 (93.6-100.0)	100.0 (93.6-100.0)		0.00	1.000 (< 0.001)	> 0.62
τ^2	100.00 (93.6-100.0)	92.86 (82.7-98.0)	14.00 (13.0-15.1)	0.00	0.992 (< 0.001)	> 0.13
R	100.00 (93.6-100.0)	100.0 (93.6-100.0)		0.00	1.000 (< 0.001)	> 1.15



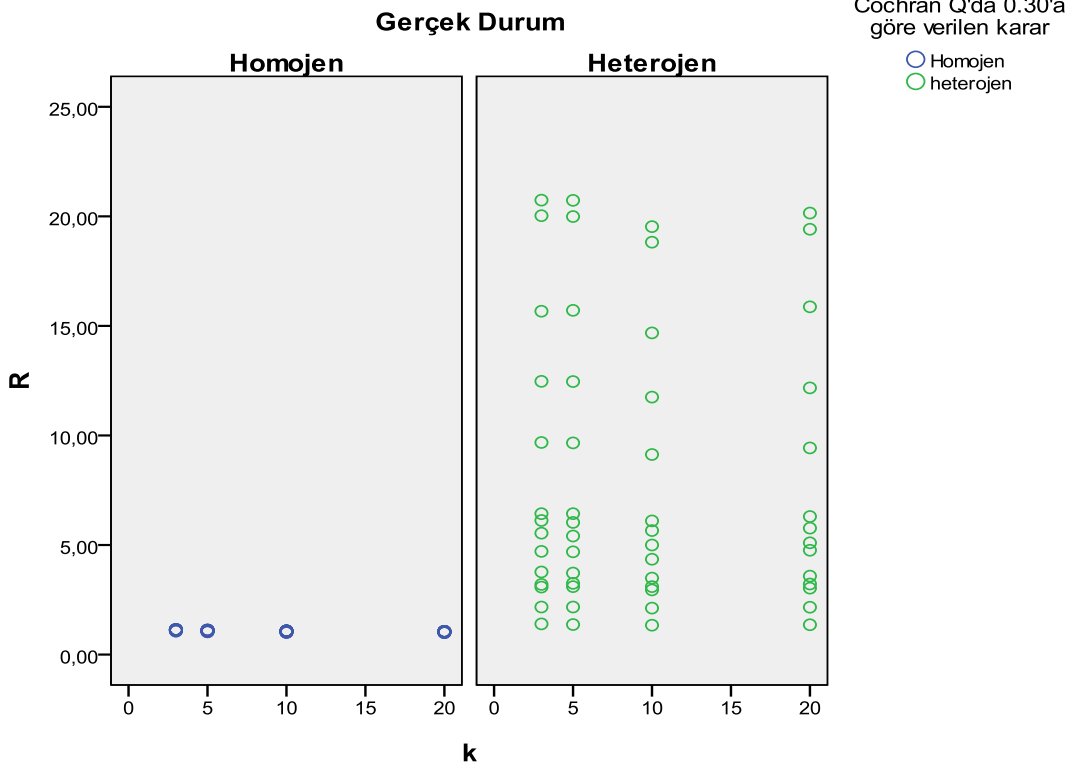
Şekil 5.17. I^2 'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



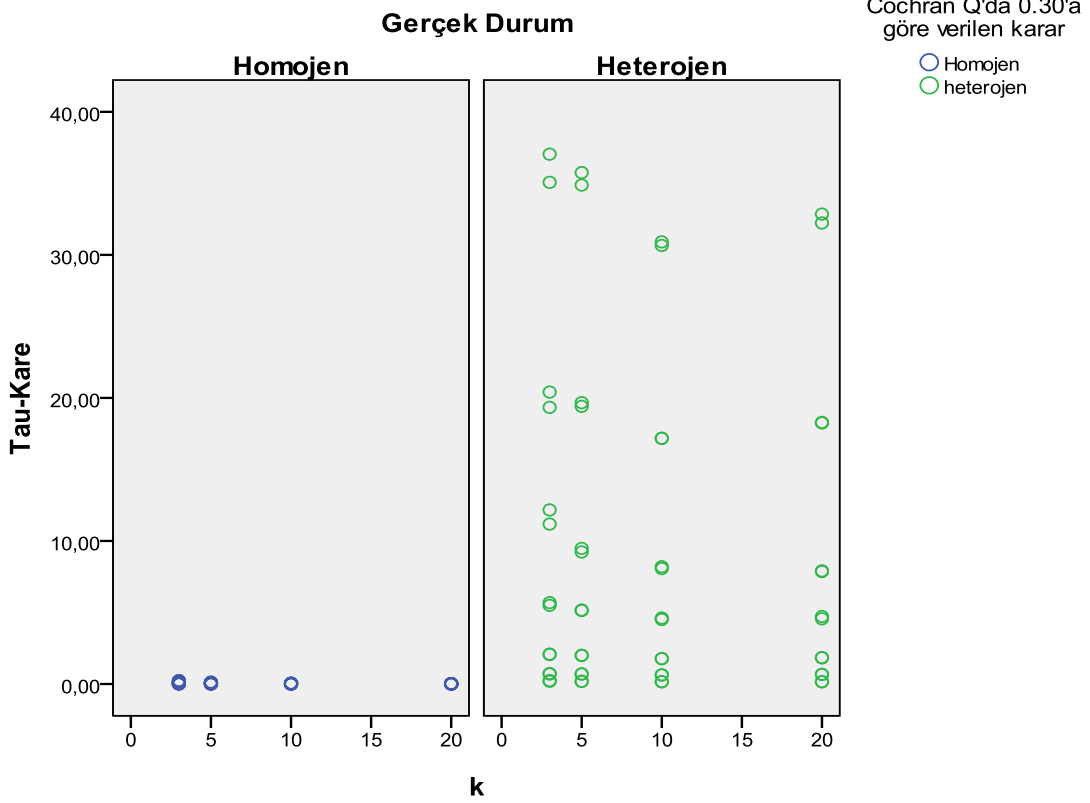
Şekil 5.18. H 'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



Şekil 5.19. H_M 'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



Şekil 5.20. R'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması



Şekil 5.21. τ^2 'nin, Cochran Q'da 0.30 anlamlılık seviyesinde verilen kararların gerçek durumla karşılaştırılması

Meta analizi çalışmalarında kullanılacak olan heterojenlik testinin ve heterojenlik ölçümlerinin I. ve II. Tip hatadan etkilenme durumlarının ortaya konması amacıyla, homojen oluşturulan durumlar için gerçekte homojen olması beklenirken heterojen olarak yorumlananlar (I. tip hata, α) ve heterojen oluşturulan durumlar için gerçekte reddedilmesi gerektiği halde homojen olarak yorumlananlar (2. tip hata β) tespit edilmiştir. Farklı tedavi oranları, çalışma sayıları, örneklem büyüklükleri ve anlamlılık seviyeleri için, 1000 denemede yapılan I. tip hata yüzdeleri Çizelge 5.5’de, yapılan 2. tip hatalardan yararlanarak testin gücü ($1-\beta$) hesaplanarak ve bunlara ait yüzdeler Çizelge 5.6’da verilmektedir.

Çizelge 5.5: 1000 denemede elde edilen Tip 1 hata yüzdeleri

	K=3						K=5					
	N=50			N=500			N=50			N=500		
$\pi_T=\pi_C$	0.05	0.10	0.3	0.05	0.10	0.3	0.05	0.10	0.3	0.05	0.10	0.3
0.05	0.5	3.2	24.1	5.4	10.6	29.2	0.4	1.6	15.1	5.7	10.8	30.9
0.10	3.3	8.5	28.3	4.5	9.4	28.2	1.7	5.9	27.3	5.6	11.9	31.3
0.20	3.8	7.8	28.6	4.6	10.2	32.4	3.7	9.1	28.6	5.0	8.7	28.1
0.50	5.3	10.0	32.2	5.2	10.2	30.0	4.1	8.1	28.1	5.6	12.3	29.9
0.80	5.1	10.5	29.2	4.6	10.1	28.3	5.1	8.9	29.4	4.4	9.3	29.9
0.90	3.8	8.1	32.2	4.3	9.7	30.4	1.5	4.3	25.5	5.1	9.5	30.7
0.95	0.4	2.5	21.8	5.7	9.7	28.7	0.3	1.3	17.4	5.0	9.7	31.6
	K=10						K=20					
	N=50			N=500			N=50			N=500		
$\pi_T=\pi_C$	0.05	0.10	0.3	0.05	0.10	0.3	0.05	0.10	0.3	0.05	0.10	0.3
0.05	0	0.2	6.3	3.6	8.4	28.5	0	0.3	2.2	5.1	8.9	30.0
0.10	1.1	2.7	18.9	4.7	9.3	30.9	0.5	2.4	17.6	4.9	9.3	29.2
0.20	4.9	9.0	29.6	5.3	10.6	31.3	4.3	6.8	26.4	5.0	10.8	32.0
0.50	4.7	9.9	26.4	3.8	8.8	29.0	4.1	7.8	27.0	6.5	9.8	29.4
0.80	2.7	6.6	25.2	5.4	10.3	29.4	3.2	7.1	26.9	5.1	10.9	32.5
0.90	1.2	4.2	22.0	4.8	10.3	30.3	0.4	1.9	16.2	4.1	8.9	27.7
0.95	0	0.3	7.9	2.7	7.3	26.8	0	0	2.4	4.5	8.3	29.1

Çizelge 5.6: 1000 denemede elde edilen Power yüzdeleri

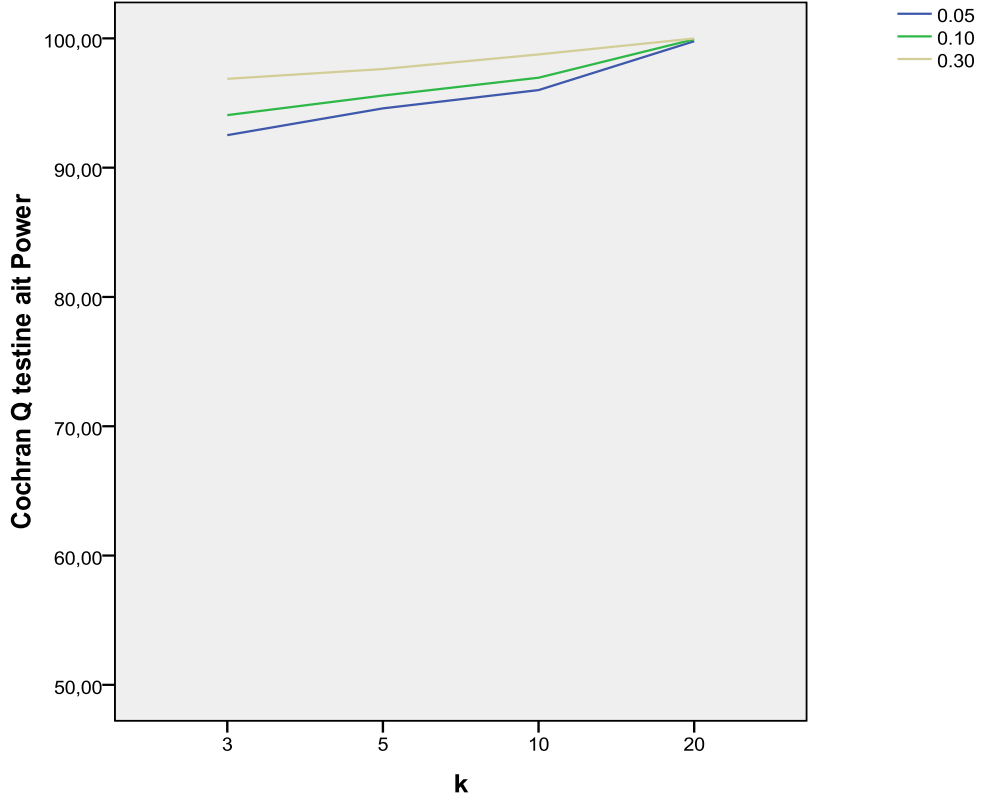
	K=3						K=5					
	N=50			N=500			N=50			N=500		
$\pi_T=\pi_C$	0.05	0.10	0.3	0.05	0.10	0.3	0.05	0.10	0.3	0.05	0.10	0.3
0.05	100	100	100	100	100	100	100	100	100	100	100	100
0.10	100	100	100	100	100	100	100	100	100	100	100	100
0.20	98.1	99.6	100	100	100	100	100	100	100	100	100	100
0.50	74.9	83.9	95.0	100	100	100	92.6	95.0	99.1	100	100	100
0.80	23.6	34.3	61.2	98.6	99.1	100	31.6	43.1	67.7	100	100	100
0.90	100	100	100	100	100	100	51.8	66.7	88.0	100	100	100
0.95	100	100	100	100	100	100	100	100	100	100	100	100

Çizelge 5. 6'nın devamı

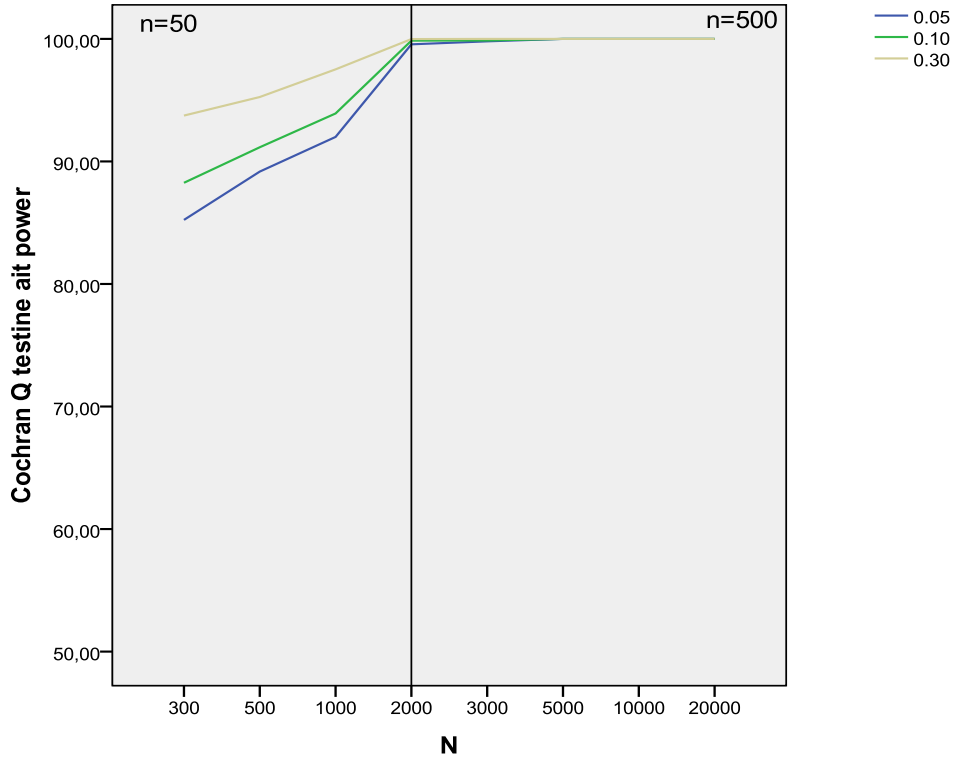
	K=10						K=20					
	N=50			N=500			N=50			N=500		
$\pi_T=\pi_C$	0.05	0.10	0.3	0.05	0.10	0.3	0.05	0.10	0.3	0.05	0.10	0.3
0.05	100	100	100	100	100	100	100	100	100	100	100	100
0.10	100	100	100	100	100	100	100	100	100	100	100	100
0.20	100	100	100	100	100	100	100	100	100	100	100	100
0.50	99.6	99.9	99.9	100	100	100	100	100	100	100	100	100
0.80	44.4	57.5	82.7	100	100	100	71.6	81.7	95.0	100	100	100
0.90	100	100	100	100	100	100	96.9	99.0	99.9	100	100	100
0.95	100	100	100	100	100	100	100	100	100	100	100	100

Heterojenlik testine ait anlamlılık seviyesi 0.05, 0.10 ve 0.30 olduğunda çalışma sayısına bağlı olarak heterojenlik testinin gücündeki değişim Şekil 5.22'de, örneklem büyüklüğüne bağlı olarak heterojenlik testinin gücündeki değişim ise Şekil 5.23'de verilmektedir. Cochran Q heterojenlik testinde anlamlılık seviyesi 0.05, 0.10 ve hatta 0.30 olarak da alınsa testin gücü % 90'ın üzerindedir. Ancak, anlamlılık seviyesi olarak 0.05 ve 0.10 alındığında meta analizine dahil edilen çalışma sayısı 10'dan büyük olduğunda testin gücü biraz daha güçlüdür. Anlamlılık seviyesi 0.30 alındığında meta analizine dahil edilen çalışma sayısı ne olursa olsun testin gücü değişmemektedir (Şekil 5.22).

Birleştirilmiş çalışmaların örneklem büyüklükleri bakımından incelendiğinde ise, meta analizine dahil edilen çalışma sayısı ne olursa olsun örneklem büyüklüğü her iki grupta da 500 olduğunda tüm anlamlılık seviyeleri için Cochran Q testinin gücü % 100 olarak tespit edilmiştir. Aynı şekilde, birleştirilmiş çalışmalarda meta analizine dahil edilen tüm çalışmalardaki toplam örneklem büyüklüğü 2000'nin üzerinde olduğu durumlarda testin gücünün % 100 olduğunu söyleyebiliriz. Tedavi ve kontrol gruplarında örneklem büyüklükleri 50 olarak alındığında ise Cochran Q testinin gücünün % 100 olarak tespit edildiği durum ancak çalışma sayısının 20 olduğu durumdur. Özellikle meta analizine dahil edilen çalışma sayısı 10'dan az olduğunda Cochran Q testinin gücü % 90'ın altına düşmektedir (Şekil 5.23).



Şekil 5.22: Farklı anlamlılık seviyelerinde ve çalışma sayılarında heterojenlik testinin gücü



Şekil 5.23: Farklı anlamlılık seviyelerinde ve örneklem büyüklüklerinde heterojenlik testinin gücü

6.SONUÇLAR VE ÖNERİLER

Meta analizi yalnızca ortalama tedavi etkilerini (analitik meta analizi) tahmin etmek için değil, aynı zamanda heterojenliğin olası kaynaklarını araştırmak için de çok önemlidir. Homojenlik analizi, etki büyüklüklerinin bir çalışmadan diğer bir çalışmaya nasıl değiştiğini gösteren bir ölçüttür. Bir meta analizinde sonuçların birleştirilmesinden sonra bu birleştirilmiş bulguları kullanmadan önce heterojenlik için istatistiksel testlerin ve grafik gösterimlerinin incelenmesi gerekmektedir. Eğer homojenlik sağlanamamışsa çalışmalar ve sonuçlar arası tutarsızlık varsa böyle bir durumda heterojenlik uygun istatistiksel yöntemlerle analiz yapılarak çalışmanın homojenliği sağlanır. Aksi takdirde bilgiler istatistiksel açıdan güvenilir olmayabilir (6,8).

Meta analizinde heterojenlik testi olarak kullanılan Cochran Q testi, meta analizine dahil edilen çalışma sayısından, örneklem büyüklüğünden ve gruplardaki tedavi etkinlik oranları arasındaki farklılıklardan etkilenmektedir. Bu amaçla heterojenlik kontrolü için heterojenlik testinin yanı sıra heterojenlik ölçümlerinin de verilmesi tavsiye edilmektedir (31,43,46,49). Bu amaçla, heterojenlik testinin ve ölçümlerinin çalışma sayısından ve örneklem büyüklüğünden etkilenme durumları araştırılmış ve simülasyon bulgularına göre, meta analizine alınacak olan çalışmaların, homojen veya heterojen olması durumlarında, heterojenlik ölçümü olarak I^2 ve τ^2 test istatistiklerinin kullanılması gerekliliği görülmüştür. Ancak bu heterojenlik ölçümlerinin de güven aralıklarına ait problemlerle karşılaşmıştır.

Buradan yola çıkarak, meta analizi çalışmalarında kullanılacak olan heterojenlik testinin ve heterojenlik ölçümlerinin I. ve II. tip hatadan etkilenme durumlarının ortaya konması ve heterojenlik ölçümlerine ait kesim noktalarının belirlenmesinin gerekliliği sonucuna varılmıştır. Cochran Q istatistiğinin dezavantajından dolayı anlamlılık seviyesi için 0.05'in yerine 0.10'un alınması ve Cochran Q heterojenlik testinin yanı sıra bir de heterojenlik ölçümlerinden olan I^2 ve H_M ölçüm değerlerinin de verilmesi tavsiye edilmektedir (43). Yani, önce heterojenlik testi ile heterojen olup olmamasına karar verilecek, eğer heterojen ise heterojenlik ölçümlerinden I^2 değerinin de %50'nin üzerinde olması durumunda çalışmaların heterojen olduğu sonucuna varılacaktır.

Elde edilen bulgularının sonucuna göre, meta analizi çalışmalarında Cochran Q testine ait anlamlılık seviyesi olarak 0.05 veya 0.10 alınıyorsa heterojenlik ölçümü olarak, gerçek homojen ve heterojenleri en iyi ayırt edebilen heterojenlik ölçümü olarak τ^2 test istatistiğinin kullanılması önerilmektedir. Anlamlılık seviyesi için 0.05 ve 0.10 alınıyorsa heterojenlik ölçümü olan τ^2 test istatistiğinin 0.24'ün üzerinde olduğu durumlarda meta analizine dahil

edilen heterojen alıřmaları tam olarak ayırabilecektir. Ayrıca, diđer heterojenlik ölçümlerinden H ve R test istatistiđi için 1.4'ün, H_M test istatistiđi için 1.1'in üzerindeki deđerler için meta analizine dahil edilen alıřmaların heterojen olduđu sonucuna varılmaktadır. Diđer bir heterojenlik ölçümü olan I² test istatistiđi için ise, Q testinin anlamlılık seviyesi 0.05 olarak alınıyorsa % 71.43'ün üzerinde, anlamlılık seviyesi olarak 0.10 alınıyorsa % 50 olarak deđil % 39.85'in üzerinde olduđunda alıřmaların heterojen olduđu sonucuna varılmaktadır. Bunun yanı sıra, Cochran Q test istatistiđi için anlamlılık seviyesi olarak 0.30 alındıđında tüm heterojenlik ölçümlerinin gerek homojen ve heterojenleri ok iyi ayırt ettiđi ortaya konulmuřtur. Heterojenliđin test edilmesinde, heterojenlik ölçümlerinden I² test istatistiđi, % 15.66'in üzerinde bir deđer aldıđında, H ve R için 1.15'in üzerinde olması durumunda, H_M için 0.62'nin üzerinde olması, alıřmalar arası varyansın da τ^2) 0.13'ün üzerinde olması durumlarında da meta analizine dahil edilen alıřmaların heterojen olduđu sonucuna varılabilmektedir.

Ayrıca, elde edilen bulgular sonucunda, meta analizine dahil edilen alıřma sayısının 20 olması durumunda, Cochran Q testine ait anlamlılık seviyesi için 0.05, 0.10 ya da 0.30 alındıđında testin gücünün % 100'lere ulařtıđını söyleyebiliriz. Bunun yanı sıra birleřtirilmiř alıřmaların toplam örneklem büyüklüđünün 1000'nin altında olduđu durumlarda tüm anlamlılık seviyesinde testin gücü % 90'nın altına düşerken, 2000'nin üzerindeki örneklem büyüklükleri için testin gücü % 100'e ulařmaktadır.

Bu alıřmada, tedavi ve kontrol gruplarında, örneklem büyüklükleri eřit olmak üzere, birleřtirilmiř alıřmalardaki örneklem büyüklükleri dikkate alınmıřtır. Gruplar arasında örneklem büyüklüklerinin dengeli olmadıđı durumlar da incelenebilir. Ayrıca, sonuç deđiřkeninin sürekli yapıda olduđu durumlar için heterojenlik testi ve heterojenlik ölçümlerinin alıřma sayısından ve örneklem büyüklüđünden etkilenme durumları incelenebilir.

7. KAYNAKLAR

1. **Normand SLT.** Tutorial in Biostatistics Meta-Analysis: Formulating, evaluating, Combining and Reporting. *Statistics in Medicine*, **1999**; 18(3): 321-359.
2. **L'abbe K.A, Detsky, AS, O'Rourke K.** Meta analysis in clinical research. *Annals of Internal Medicine*, **1987**;107(2): 224-233.
3. **Gliner JA, Morgan GA, Harmon RJ.** Meta-Analysis: Formulation and Interpretation. *Journal of the American Academy of Child & Adolescent Psychiatry*, **2003**; 42(11): 1376-1379.
4. **Boissel JP, Blanchard J, Panak E.** Considerations for the Meta-Analysis of Randomized Clinical Trials. *Controlled Clinical Trials*, 1989; 10 (3): 254-281.
5. **Didem D.** Klinik çalışmalarda meta analizi uygulamaları. Yayınlanmamış yüksek lisans tezi, Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun, **2005**.
6. **Lipsey MV, Wilson DB.** Practical meta-analysis. Sage Publication. Thousand Oaks, CA. **2001**.
7. **Elwood M.** Critical Appraisal of Epidemiological Studies and Clinical Trials. 3rd Ed. NewYork: Oxford Medical Pub.; **2003**; 198-217.
8. **Bailer III JC, Mosteller F.** Reviews and meta-studies, Medical Uses of Statistics. 2nd ed. Boston, NEJM Books,**1992**.
9. Örer H., Meta Analizlerin Rolü. Erişim Adresi: www.toraks.org.tr/mse2-ppt-pdf/10_hakan_orer_word.pdf. Erişim tarihi: 7.01.2009.
10. **Egger M, Smith GD, Altman DG.** Systematic Reviews in Health Care: meta analysis in Context. 2nd ed. London, BMJ Books: **2001**.
11. **Davey-Smith G, Egger, M, Phillips AN.** Meta-analysis: beyond the Grand mean? *British Medical Journal*, **1997**; 315: 1610-1614..
12. **Deeks JJ.** Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine* **2002**; 21: 1575-1600.
13. **Anderson P, Green S.** Cochrane Collaboration open learning material for reviewer. Version 1.1.The Cochrane Collaboration. **2002**.
14. **Sutton, AJ, Abrams KR, Jones DR, Sheldon TA, Song F.** Methods for Meta-analysis in Medical Research, Third Avenue, New York, John Wiley & Sons, Ltd, **2000**.
15. **Borenstein M, Hedges L, Rothstein H.** Meta-Analysis Fixed effect vs. random effects, Erişim Adresi: www.Meta-Analysis.com, 2007.
16. **Schechtman E.** Odds Ratio, Relative Risk, Absolute Risk Reduction and the Number Needed to Treat- Which of These Should We Use? *Value in Health*, **2002**; 5(5): 431-436.
17. **Altman DG.** Confidence intervals for the number needed to treat. *BMJ*, **1998**; 317: 1309-1312.
18. **Süt N, Şenocak M.** Relatif risk ölçütünün odds oranı, atfedilen risk ve tedaviye gerekli sayı ölçütleriyle karşılaştırılması. *Trakya Üniversitesi Tıp Fakültesi Dergisi* **2007**; 24(3): 213-221.
19. **Newcombe RG.** Confidence intervals for the number needed to treat. Absolute risk reduction is less likely to be misunderstood. *BMJ* **1999**; 318 (700): 1765-1767.

20. **Rosenthal R, DiMatteo MR.** Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annu. Rev. Psychol.*, **2001**; 52: 59-82.
21. **Borenstein M, Hedges LV, Higgins JPT, Rothstein HR.** Introduction to Meta Analysis. 1st ed., United Kingdom, John Wiley & Sons, Ltd, **2009**.
22. **Rosenberg MS, Adams DC, Gurevitch J.** MetaWin Statistical Software for Meta-Analysis Version 2.0, Sinauer Associates, Inc, **2000**, page: 13-16.
23. **Furr M.** Psychology Department, Wake Forest University, Summary of Effect Sizes and Their Links to Inferential Statistics; **2008**.
24. **Fern EF., Monroe KB.** Effect-size estimates: Issues and problems in Interpretation. *The Journal of Consumer Research*, **2009**; 23(2)(Sep, 1996): 89-105.
25. **DeCoster J.** Meta-analysis notes. Institute for Social Science Research University of Alabama, **2009**.
26. Wikipedia, The Free Encyclopedia, Effect Size, Eriřim Adresi: http://en.wikipedia.org/wiki/Effect_size; Eriřim Tarihi: 23.06.2009.
27. **Hartung J, Knapp G, Sinha BK.** Statistical Meta-Analysis With Applications. New Jersey, John Wiley & Sons., **2008**.
28. Eriřim Adresi: www.Meta-Analysis.com, Eriřim Tarihi:12.04.2009.
29. **Ledesma RD, Macbeth G, Kohan NC.** Computing effect size measures with VİSta- The Visual Statistics System. *Tutorial in Quantitative Methods for Psychology*, **2009**; 5(1): 25-34.
30. **Durlak JA.** How to Select, Calculate and Interpret Effect Size. *Journal of Pediatric Psychology*, **2009**; 1-12.
31. **Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J.** Assessing heterogeneity in meta-analysis: Q statistics or I² Index?. *Psychological Methods*, **2006**; 11(2):193-2006.
32. Wikipedia, The Free Encyclopedia, Mean Difference, Eriřim Adresi: http://en.wikipedia.org/wiki/Mean_difference, Eriřim Tarihi: 23.06.2009.
33. **Friedrich JO, Adhikari NKJ, Beyene J.** The ratio of means methods as an alternative to mean difference for analyzing continuous outcome variables in meta-analysis:A simulation study. *BMC Medical Research Methodology*, **2008**;8:32.
34. **McGraw K, Wong S.** A common language effect size statistics. *Psychological Bulletin*, **1992**; 111(2): 361-365.
35. **Cooper H, Hedges LV.** The Handbook of Research Synthesis. New York: Russell Sage Foundation, **1994**.
36. **Bouzas JC, Takkouche B, Cadarso-Suarez C, Spiegelman D.** HEpiMA: software for the identification of heterogeneity in meta-analysis. *Computer Methods and Programs in Biomedicine* **2001**; 64(2), 101-107.
37. Bland M. Meta-analysis: dealing with heterogeneity, 2006. Eriřim Adresi: http://www-users.york.ac.uk/~mb55/msc/systrev/week7/het_text.pdf. Eriřim Tarihi: 10.03.2009.
38. Evidence Based Neurology BMJ Books 2007; Eriřim Adresi: http://www.poems.msu.edu/EBM/Overviews/steps_in_a_metaanalysis.htm; Eriřim Tarihi: 16.06.2009.
39. Steps in a meta-analysis. Eriřim tarihi: 16.06.2009. Eriřim Adresi: <http://pages.infinit.net/rlevesqu/Syntax/MetaAnalysis/MetaAnalysisOfContOutcomes.txt>.

40. **Xu H, Platt RW, Luo ZC, William S, Fraser WD.** Exploring heterogeneity in meta analyses: needs, resources and challenges. *Paediatric and Perinatal Epidemiology*, **2008**; 22 (Suppl.1): 18-28.
41. **Higgins JPT, Green S.** Cochrane Handbook for Systematic Review of Interventions 4.2.6., The Cochrane Collobariton, **2006**.
42. **Abramson, JH,** Making Sense of Data. 2nd Ed. New York: Oxford University Press, **1994**.
43. **Mittlböck M, Heinzl H.** A simulation study comparing properties of heterogeneity measures in meta analyses. *Statistics in Medicine*, **2006**; 25(24): 4321-4333.
44. **Cochran WG.** The combination of estimates from different experiments. *Biometrics* **1954**; 10 (1):101–129.
45. **Hardy RJ., Thompson SG.** Detecting and Describing Heterogeneity in Meta Analysis, *Statistics in Medicine*, **1998**; 17(8): 841-856.
46. **Rücker G, Schwarzer G, Carpenter JR, Schumacher M.** Undue reliance on I^2 in assessing heterogeneity may mislead, *BMC Medical Research Methodology*, **2008**; 8: 79.
47. **Takkouche B, Cadarso-Suarez C, Spiegelman D.** Evaluation of old and new tests of heterogeneity in epidemiologic meta-analysis. *American Journal of Epidemiology*, **1999**; 150 (2): 206-215.
48. **Baujat B, Mah C, Pignon J, Hill C.** A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials. *Statistics in Medicine*, **2002**; 21(18):2641–2652.
49. **Virgili G, Conti AA, Moja L, Gensini GF, Gusino R.** Heterogeneity and meta analyses: do study results truly differ? *Internal and Emergency Medicine*, **2009**; 4:423-427.
50. **Higgins JPT, Thompson SG.** Quantifying heterogeneity in a meta analysis. *Statistics in Medicine*, **2002**; 21 (11):1539-1558.
51. **Patsopoulos N, Evangelou E., Ioannidis JPA.** Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. *International journal of Epidemiology*, **2008**; 37 (5): 1148-1157.
52. **Song F, Sheldon TA, Sutton AJ, Abrams KR, Jones DR.** Methods for Exploring Heterogeneity in Meta-analysis. *Evaluation & The Health Professions*. 2001; 24(2):126-151.
53. **Hedges LV, Olkin I.** Statistics Methods for Meta-analysis, Academic Press, Boston, **1985**.
54. **Bax L, Ikeda N, Fukui N, Yaju Y, Tsuruta H, Moons KGM.** More than numbers: The Power of Graphs in Meta Analysis. *American Journal of Epidemiology*. **2009**; 169 (2):249-255.
55. **Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L.** Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine*, **2007**; 26(25):4544-4562.
56. **Wang MC, Bushman BJ.** Using the Normal Quantile Plot to Explore Meta-Analytic Data Sets. *Psychological Methods*, **1998**; 3(1):46-54.
57. **Bailey, KR.** Inter-study differences: How should they influence the interpretation and analysis of results? *Statistics in Medicine*, **1987**; 6:351-358.
58. **Leandro G.** Meta-Analysis in Medical Research - The Handbook for the Understanding and Practice of Meta-Analysis, *Blackwell Publishing*, **2005**; ISBN 1-4051-2733-3, page:39-42.
59. **Whitehead A.** Meta analysis of controlled clinical trials, 1st Ed. England, John Wiley & Sons. **2002**
60. MINITAB Inc. Version: 15 Trial Software. **2007**.

61. **Erdoğan S, Kanık EA.** Meta Analizine Giren Çalışmaların Homojen Olduğu Durumlarda Binary Etki Büyüklüklerinin Örneklem Büyüklüğünden ve Tedavi Etkinliği Oranlarından Etkilenme Durumları: Bir Ön Çalışma, *Türkiye Klinikleri Biyoistatistik Dergisi*, Kabul edildi.
62. **Sanchez-Meca J, Marin-Martinez F.** Testing the significance of a common risk difference in meta analysis. *Computational Statistics&Data Analysis*, **2000**; 33 (3): 299-313.
63. StatSoft. Inc. STATISTICA System Reference. Version 6.0, **2002**.
64. IBM SPSS Statistics. Version 19.0 Trial Software, Chicago: SPSS Inc., **2010**.
65. **Petitti DB.** Approaches to heterogeneity in meta-analysis. *Statistics in Medicine*, **2001**; 20 (23):3625–3633.

ÖZGEÇMİŞ

1972 yılında Bayburt'ta doğdu. İlk, orta ve lise öğrenimini Zonguldak'ta tamamladı. 1988 yılında girdiği Hacettepe Üniversitesi Fen Fakültesi İstatistik bölümünden 1992 yılında mezun oldu. 1994 yılında Mersin Üniversitesi Tarsus Meslek Yüksekokulunda Öğretim Görevlisi olarak göreve başladı. 2002 yılında Mersin Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim dalında Yüksek Lisans, 2005 yılında da Doktora öğrenimine başladı. 2003 yılından itibaren de Mersin Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim dalında öğretim görevlisi olarak çalışmaktadır.