Türkiye Klinikleri Biyoistatistik Dergisi
Turkiye Klinikleri Journal of Biostatistics Turkiye Klinikleri J Biostat. 2023;15(1):67-78

ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH DOI: 10.5336/biostatic.2022-91725

Odds Oranı için Logit Güven Aralıkları: Bir Benzetim Çalışması

Logit Confidence Intervals for Odds Ratio: A Simulation Study

İsmet DOĞAN[a], Nurhan DOĞAN[a]

[a]Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim ABD, Afyonkarahisar, Türkiye

ÖZET Amaç: Bu makalenin amacı, odds oranı için hesaplanan Woolf logit, Gart düzeltilmiş logit ve Agresti bağımsız-düzleştirilmiş logit güven aralıklarını farklı örneklem büyüklükleri için karşılaştırmaktır. **Gereç ve Yöntemler:** Çalışmada Phyton-random kütüphanesi kullanılarak $10 \leq n \leq 1000$ aralığında yer alan 35 farklı n değeri için veri türetilmiştir. Verilerin türetilmesinde önce n_{11}, n_{12}, n_{21} ve n_{22} ile gösterilen frekanslardan hangisine değer atanacağı sonra da ilgili frekans değeri belirlenmiştir. $n = 10$ için 286, $n = 15$ için 815 ve $n \geq 20$ için 1000'er farklı veri seti çalışmada kullanılmıştır. Ancak odds oranları ile güven aralığı hesaplamalarında n_{11}, n_{12}, n_{21} ve n_{22} frekanslarından herhangi birinin sıfır olması durumunda hesaplama yapılamadığı için simülasyon sayıları her bir örneklem büyüklüğü ve yöntem için farklılık göstermiştir. **Bulgular:** n_{11}, n_{12}, n_{21} ve n_{22} frekanslarından herhangi birinin değerinin diğerlerine göre aşırı derecede farklılaşması odds oranının değerini önemli derecede etkilemekte, güven aralığını ise büyütmektedir. Odds değerleri ve güven aralığı açısından bakıldığında yöntemlerden elde edilen değerler $Agresti \geq Gart \geq Woolf$ şeklinde sıralanmaktadır. Özellikle $n \geq 150$ olması durumunda her üç yöntemden elde edilen güven aralıkları istatistiksel olarak anlamlı hale getirmektedir. **Sonuç:** Çalışmada dikkate alınan güven aralıkları ile ilgili yöntemlerden hiçbirinin en iyi yöntem olarak belirtilemeyeceği, bu yöntemlerden yararlanılarak elde edilen güven aralıklarının çok geniş ve nokta tahminleri konusunda simetrik olmadıkları belirlenmiştir. Çalışmada dikkate alınan tüm örneklem büyüklükleri için Woolf yöntemi, hesaplanan odds oranının istatistiksel olarak anlamlı olup olmaması, güven aralığının büyüklüğü ve $alt\ sınır < odds\ oranı < üst\ sınır$ kriterleri dikkate alındığında diğer iki yönteme göre daha iyi performans göstermiştir. Ancak n_{11}, n_{12}, n_{21} ve n_{22} frekanslarından herhangi birinin sıfır olması durumunda Woolf yönteminden elde edilecek güven aralığının anlamsız olduğu göz ardı edilmemelidir.

Anahtar kelimeler: Odds oranı; Woolf logit güven aralığı;
Gart düzeltilmiş logit güven aralığı;
Agresti bağımsız-düzeltilmiş logit güven aralığı;
iki sonuçlu veri

ABSTRACT Objective: The aim of this article is to compare the Woolf logit, Gart adjusted logit and Agresti independent-smoothed logit confidence intervals calculated for odds ratio for different sample sizes. **Material and Methods:** In the study, data were derived for 35 different n values in the range of $10 \leq n \leq 1000$ using the Phyton random library. In the derivation of the data, first which of the frequencies indicated by n_{11}, n_{12}, n_{21} and n_{22} will be assigned a value, then the relevant frequency value is determined. 286 different data sets for $n = 10$, 815 for $n = 15$ and 1000 different data sets for $n \geq 20$ were used in the study. However, in the calculations of odds ratios and confidence intervals, if any of the n_{11}, n_{12}, n_{21} and n_{22} frequencies is zero, the calculation cannot be made, so the simulation numbers are different for each sample size and method. **Results:** Excessive variation in the value of any of the n_{11}, n_{12}, n_{21} and n_{22} frequencies from the others significantly affects the value of the OR and enlarges the confidence interval. In terms of odds values and confidence intervals, the values obtained from the methods are listed as $Agresti \geq Gart \geq Woolf$. Especially in the case of $n \geq 150$, the confidence intervals obtained from all three methods make it statistically significant. **Conclusion:** It has been determined that none of the methods related to the confidence intervals considered in the study can be specified as the best method, and the confidence intervals obtained by using these methods are very wide and not symmetrical in point estimations. For all sample sizes considered in the study, Woolf's method performed better than the other two methods, considering whether the calculated odds ratio was statistically significant, the size of the confidence interval, and the $lower\ limit < odds\ ratio < upper\ limit$ criteria. However, if any of the n_{11}, n_{12}, n_{21} and n_{22} frequencies is zero, it should not be ignored that the confidence interval to be obtained from the Woolf method is meaningless.

Keywords: Odds ratio; Woolf logit interval;
Gart adjusted logit interval;
Agresti independence-smoothed logit interval;
binary data

Correspondence: Nurhan DOĞAN
Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim ABD, Afyonkarahisar, Türkiye
E-mail: nurhandogan@hotmail.com

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

Received: 01 Jun 2022 ***Received in revised form:*** 06 Sep 2022 ***Accepted:*** 06 Oct 2022 ***Available online:*** 12 Dec 2022

Tıbbi araştırmalarda genellikle iki sonuçlu veriler kullanılarak 2 farklı tedavinin karşılaştırılmasında yaygın olarak risk farkı (2 oranın farkı), rölatif risk [relative risk (RR)] veya odds oranı [odds ratio (OR)] olarak ifade edilen parametreler kullanılmaktadır. Risk farkı, etkinin mutlak bir ölçümüdür, RR ve OR ise sonuçları karşılaştırmak için göreceli ölçümlerdir. Geriye dönük vaka-kontrol çalışmalarında, risk farkı ve RR tahmin edilemediği için OR kullanılır. Lojistik regresyonda, OR'nin regresyon katsayısı ile doğrudan bir ilişkisi olduğu da iyi bilinmektedir. RR, randomize kontrollü çalışmalarda ve kohort çalışmalarında, özellikle 2 ilgili oranın her ikisi de küçük olduğunda kullanılır. Böyle bir durumda risk farkı, RR kadar bilgilendirici değildir. RR ve OR, hastalık nadir olduğunda karşılaştırılabilir. Bazı yaygın hastalıklar için OR'nin değeri fazla tahmin edilebilir. Bu durumda, OR'nin yerine RR kullanılmalıdır.[1] Doğru kullanıldıklarında OR'ler, bir grubun diğerine karşı taşıdığı riskteki değişikliğin tahminine izin veren güçlü bir istatistiktir. Bir grubun diğer bir gruba karşı üstlendiği riskin kaç kat olduğunu ve nedensellik varsayımıyla, bir müdahalenin sonuçlar üzerinde bir etkisinin olup olmadığını hızlı bir şekilde gösterir.[2] OR bir sonuç açısından 2 grubu (tedaviler ya da risk faktörleri) karşılaştıran etki büyüklüğünün değerlendirilmesinde kullanılan yaygın bir ölçüdür, ancak OR tam olarak anlaşılamamıştır. OR'yi sadece orantısal bir değişiklik olarak tanımlamak doğru değildir. RR daha sezgiseldir, ancak vaka kontrol çalışmalarından veya nadir durumlar dışında lojistik regresyonlardan elde edilemez. Ayrıca OR'ler bildirilirken yanlış yapma tehlikesi de bulunmaktadır. Sonuç nadir olduğunda, OR ve RR yaklaşık olarak aynıdır. Bilimsel çalışmalarda bazen bu yaklaşıma çok fazla güvenilerek, OR riskler açısından tartışılmaktadır. OR, çalışmada "nadir" olmayan sonuçlar için zayıf bir yaklaşım olmakla kalmaz, aynı zamanda yanlızca OR'nin verildiği bir istatistiksel analizde OR, RR anlamına da gelmez. Ayrıca RR aynı zamanda başlangıç veya kontrol grubundaki sonucun riskine de bağlıdır.[3] OR özellikle tıbbi araştırmalarda yaygın olarak kullanılmaktadır. Çünkü OR;

- İki sonuçlu 2 değişken arasındaki ilişki için bir tahmin (güven aralığı ile) sağlar,
- Lojistik regresyon kullanarak diğer değişkenlerin bu ilişki üzerindeki etkilerinin incelenmesini sağlar,
- Vaka kontrol çalışmalarında özel ve çok uygun bir yoruma sahiptir.[4]

OR aynı zamanda bir olayın gerçekleşmesinin bir faktöre maruz kalma ile ne kadar güçlü bir şekilde ilişkili olduğunun da bir ölçüsüdür. OR'nin büyüklüğüne "ilişkinin gücü" denir. OR'nin değeri 1'den ne kadar büyükse, maruz kalma ile ilgilenilen olay arasındaki ilişkinin nedensel olması o kadar olasıdır. Maruz kalmanın belirli bir olaya yol açma olasılığının belirlenmesine yardımcı olan OR ne kadar büyük olursa, olayın maruz kalma ile ortaya çıkma olasılığı o kadar yüksek olur. OR epidemiyolojide sıklıkla kullanılır ve vaka kontrol çalışmalarında doğal ölçüdür çünkü bu çalışmalarda, oranlar arası farkın ve RR'nin kullanılması mümkün değildir. Genel olarak ifade edilecek olursa OR kesitsel, ileriye dönük ve geriye dönük çalışmalarda yaygın olarak kullanılmaktadır. Yeni tedavileri değerlendirmeye yönelik eş değerlik çalışmaları, OR ile ilgili çıkarımlarda büyük ilgi uyandırmaktadır.[5] Bir çalışmadan elde edilen OR'nin anlamlı olup olmadığını söyleyebilmek için güven aralığı hesaplanır. Araştırmacılar, örneklemden elde edilen sonuçlardan örneklemin temsil ettiği popülasyona ilişkin çıkarımlarda bulunmak isterler ve güven aralıklarının hesaplanması, bu tür çıkarımların yapılmasında son derece yararlıdır. İstatistiksel terminoloji bağlamında istatistikler, gerçek popülasyon parametresinin en olası tahminini temsil etmelerine rağmen bunlar kesin olmayan tahminlerdir. Güven aralığı, gerçek popülasyon değerinin içinde yer alması muhtemel olan bir dizi değer sağladığından bu belirsizliği ölçmenin bir yoludur. Güven aralığı, *p* değerinden daha fazla bilgi verir. Araştırmacının bir riskin/farkın büyüklüğünü karşılaştırmasını sağlar. Bu önemlidir, çünkü 2 popülasyon veya 2 tedavi grubu arasındaki risk/fark istatistiksel olarak anlamlı olabilir ancak klinik olarak anlamlı olmayabilir. Güven aralığı ayrıca istatistiksel anlamlılığın veya anlamsızlığın örneklem büyüklüğünün bir fonksiyonu olup olmayacağına dair bir gösterge olarak da kullanılabilir. Aralığın genişliği, tahmindeki değişkenlik miktarını gösterir ve tahminin alındığı örneklem büyüklüğünün bir yansımasıdır. Genel bir kural olarak, örneklem büyüklüğü ne kadar büyük olursa, değişkenlik o kadar az ve aralık o kadar dar olur, bu da daha kesin bir tahmine yol açar. Tersine daha küçük örneklem büyüklükleri için aralık genellikle daha geniştir, bu da daha fazla değişkenlik ve daha az kesinlik göstergesidir.[6] Bilimsel çalışmalarda sıklıkla kullanılan OR'ye ait güven aralıkları ile ilgili yöntemler logit, kesin (exact) ve mid-p başlıkları altında toplanmaktadır.[7] Hesaplama kolaylığından dolayı logit güven aralıkları kesin (exact) ve mid-p güven aralıklarına tercih edilmektedir. Popülasyona ait OR örneklemden elde edilen OR'den farklı olabilir. Güven aralığı, popülasyona ait gerçek OR için değerin içinde yer alması beklenen bir aralık verir. Güven aralığını hesaplamak için önem düzeyi/alfa belirtilir. Alfa değerinin 0,05 alınması, güven aralığının %95 güvenirlikte olduğu anlamına gelir. Tıp literatüründe geleneksel olarak %95'lik bir güven seçilir ancak başka güven aralıkları da kullanılabilir.[8] Güven aralığı, popülasyondan örnekleme değil örneklemden popülasyona doğru hareket eden bir çıkarım yöntemidir. Dolayısıyla güven aralığı:

- Çalışma öncesinde bir yokluk hipotezi kurulmasını gerektirmez.
- Güven aralığı, muhtemel değerlerden oluştuğu için araştırmalarda elde edilen sonucun mutlak olarak algılanmasını önler. Bir araştırmada elde edilen sonucun, elde edilebileceklerden biri olduğunun anlaşılmasını kolaylaştırır.
- Güven aralığının genişliği, çalışmanın tutarlığı, diğer bir deyişle, çalışmanın yinelenmesi hâlinde benzer sonuçlar elde edilme ihtimali hakkında fikir verir.
- Metaanaliz çalışmalarında, farklı çalışmalardan elde edilen bilgilerin birleştirilmesini kolaylaştırır.[9]

OR 1'den büyükse, maruz kalan grupta risk daha fazladır. OR 1'den küçükse, maruz kalan grupta risk daha azdır. OR 1'e eşit ise risk her iki grupta da aynıdır. OR ne kadar büyük olursa, maruz kalma ve ilgilenilen olay/hastalık arasındaki ilişki o kadar güçlü olur. Tersine OR'nin değeri 1'den küçük olduğunda, maruz kalan grupta ilgilenilen olayın/hastalığın meydana gelme olasılığı daha düşüktür. Birden küçük OR bazen koruyucu etki olarak adlandırılır. Güven aralığının 1'i içerdiği durumda maruz kalmanın, belirtilen güven düzeyinde olayın gerçekleşme olasılığını artırıp artırmadığı belirsizdir. OR'nin yorumlanmasından da anlaşılacağı üzere eğer güven aralığı 1 değerini içeriyorsa, hesaplanan OR istatistiksel olarak anlamlı kabul edilmeyecektir.[10] Üstelik OR için elde edilen güven aralığının 1 değerini içermesi, yetersiz bir örneklem büyüklüğünün de açık bir göstergesidir.[11] Normal dağılım gösteren sürekli değişkenli durumdan farklı olarak, OR'ye ait güven aralığının, nokta tahminleri konusunda simetrik olmadığı söylenmektedir. Bu çıkarım, geleneksel toplamsal yapıya dayandığından, OR yerine, ln OR'nin simetrik olduğu gösterilmiştir. Log ölçeğinin sezgisel olmadığı konusunda fikir birliği vardır ve bu nedenle ln OR'nin güven aralığı daha sonra doğal ölçekte antilog alınarak ifade edilir. Ancak ln OR'den OR'ye dönüşümden sonra OR'ye ait güven aralığının nokta tahminiyle ilgili simetriye sahip olmadığı ve bu nedenle OR'nin güven aralığının yapısal olarak sezgisel olmadığı ifade edilmektedir. Ancak OR'ye ait güven aralığının, OR'nin nokta tahmini için simetrik olduğu, sadece çarpımsal yapıda olduğu matematiksel olarak gösterilmiştir.[12] Çalışmada OR için hesaplanan Woolf logit, Gart düzeltilmiş logit ve Agresti bağımsız-düzleştirilmiş logit güven aralıklarının farklı örneklem büyüklükleri için OR'nin istatistiksel olarak anlamlı olup olmaması ($alt\ sınır < 1 < üst\ sınır$), güven aralığının büyüklüğü ve $alt\ sınır < OR < üst\ sınır$ kriterleri dikkate alınarak karşılaştırılması amaçlanmıştır.

GEREÇ VE YÖNTEMLER

Çalışmada, 2×2 tablolardan elde edilen OR ile ilgili Woolf logit, Gart düzeltilmiş logit ve Agresti bağımsız-düzleştirilmiş logit güven aralıklarını karşılaştırmak için bir simülasyon çalışması yapılmıştır. Simülasyon, benzetimi yapılan gerçek sistemin çeşitli senaryolardaki muhtemel performans göstergelerinin yüksek güven seviyesinde ve kısa bir sürede doğru olarak tahmini için yaygın olarak kullanılan ve genel kabul görmüş bir tekniktir. Herhangi bir sistemin/yöntemin davranışının incelenmesini, farklı parametrelerin çalışma durumuna etkilerinin araştırılmasını, değişen koşullar altında sistemin davranışlarının tahmin edilmesini ve gerekli önlemlerin alınmasını sağlar. Çalışmada, Phyton-random kütüphanesi kullanılarak $10 \leq n \leq 1.000$ aralığında yer alan 35 farklı n değeri (10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1.000) için veri türetilmiştir.

Veri türetimine ilişkin Python 3.9.10 programlama dili (Python Software Foundation, USA) kullanılarak yazılan programda yer alan simülasyon kurgusunun detayları aşağıdaki gibidir:

Adım 1. Veri türetimi için gerekli değerler (toplam kaç adet veri türetilmek istendiği, kategorilere atanacak değerlerin toplamı, kategori sayısı) programa girilir.

Adım 2. Python random kütüphanesindeki choice fonksiyonu kullanılarak rastgele bir şekilde kategori seçimi yapılır.

Adım 3. Choice fonksiyonu tekrar kullanılarak seçilen kategori için 0 ile istenen toplam arasında rastgele bir tam sayı seçimi yapılır. Örneğin istenen toplam 30 ise [0, 30] aralığında rastgele bir tam sayı seçilir.

Adım 4. Tüm kategoriler için 2 ve 3. adımlar tekrarlanır ($k'inci$ kategori için bir tam sayı seçildikten sonra seçilen sayıların toplamına bakılır. Eğer bu toplam, istenilen toplamdan fazla ise bu geçersiz bir veri olacağından işleme baştan başlanır. Son kategoriye atanacak tam sayı ise istenilen toplamdan mevcut toplamın çıkarılmasıyla bulunur.

Adım 5. Türetilen verinin daha önce türetilip türetilmediğine bakılır. Aynı veri daha önce türetilmişse veri setine eklenmez.

Adım 6. İstenilen toplam veri sayısına ulaşana kadar 1. adımdan itibaren program bir döngü içinde tekrar çalıştırılır. (Bu çalışmada standardizasyonu sağlamak amacıyla her bir örneklem genişliği için 1.000 adet veri türetilmesi planlanmış ancak 4 kategori için toplamı 10 veya 15 olan ve tekrar etmeyen 1.000 tane veri bulmak imkânsız olduğundan $n = 10\ için\ 216\ ve\ n = 15\ için\ 815$ adet veri türetilmiştir.)

Adım 7. Son olarak türetilen verilerden OR ve güven aralıkları hesaplanır.

$n = 10$ için 286, $n = 15$ için 815 ve $n \geq 20$ için biner farklı veri seti çalışmada kullanılmıştır. Ancak OR ile ilgili güven aralığı hesaplamalarında $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarından herhangi birinin 0 olması durumunda hesaplama yapılamadığı için veri seti sayıları her bir örneklem büyüklüğü ve yöntem için farklılık göstermiştir. Farklı örneklem büyüklükleri ve yöntemler için hesaplamalarda kullanılan simülasyon sayıları Tablo 1'de verilmiştir.

OR'nin hesaplanmasında kullanılan formüller dikkate alındığında klasik 2×2 tablo örneği Tablo 2'de verildiği şekilde oluşturulmaktadır.

OR'nin değeri $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarına bağlıdır. $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$'ye dayalı OR Eşitlik 1 kullanılarak elde edilmektedir.

$$OR = \frac{n_{11}n_{22}}{n_{12}n_{21}} \qquad (1)$$

Eşitlik 1'den de görüleceği üzere OR, 2 odds'un birbirine oranıdır. Bir olayın olasılığı her zaman pozitif olduğundan, OR her zaman pozitiftir ve değeri 0 ile çok büyük değerler arasında değişir. OR'nin 1 veya 1'e yakın değer alması, ilgilenilen olayın maruziyet ile ilişkili olmadığını, 1'den büyük olması, ilgilenilen olay için maruziyetin bir risk faktörü olabileceğini, 1'den küçük olması ise ilgilenilen olayın maruziyet sonucu gerçekleşme olasılığının daha az olduğunu ifade etmektedir. $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarına sahip 2×2'lik bir tabloda, OR maruz kalma grubundaki olayın oranının (n_{11}/n_{12}), maruz kalmayan gruptaki olayın oranına (n_{21}/n_{22}) bölümüdür. Dolayısıyla OR $(n_{11}/n_{12})/(n_{21}/n_{22})$ olur ve bu da Eşitlik 1'de verilen hâlde basitleştirilir. İlk olarak Woolf tarafından önerilen logit güven aralığı, OR tahmininin logaritmasının yaklaşık normal dağılımına dayanmaktadır. OR'nin güven aralığı;

$$e^{\ln(\widehat{OR}) \mp z_{\alpha/2} * \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}} \qquad (2)$$

eşitliği ile hesaplanmaktadır.[13] $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarından herhangi biri 0 ise güven aralığı anlamsız hâle gelir. Eşitlikte α anlamlılık düzeyini göstermektedir. Gart düzeltilmiş logit güven aralığını hesaplamak için $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarının yerine,

TABLO 1: Simülasyon sayıları.

n	Woolf logit yöntemi	Gart düzeltilmiş logit yöntemi	Agresti bağımsız-düzeltilmiş logit yöntemi
10	165	286	246
15	560	815	756
20	754	1.000	958
25	782	1.000	963
30	836	1.000	980
35	853	1.000	986
40	873	1.000	986
45	878	1.000	992
50	891	1.000	995
55	908	1.000	994
60	911	1.000	993
65	920	1.000	997
70	937	1.000	996
75	918	1.000	996
80	938	1.000	998
85	929	1.000	999
90	935	1.000	999
95	937	1.000	996
100	945	1.000	996
125	956	1.000	999
150	962	1.000	1.000
175	973	1.000	1.000
200	967	1.000	1.000
225	969	1.000	999
250	979	1.000	1.000
300	985	1.000	1.000
350	986	1.000	1.000
400	978	1.000	1.000
450	985	1.000	1.000
500	984	1.000	1.000
600	994	1.000	1.000
700	983	1.000	1.000
800	993	1.000	1.000
900	994	1.000	1.000
1.000	992	1.000	1.000
Toplam	31.550	34.101	33.824

TABLO 2: Odds oranı hesaplamaları için 2x2 tablo.

	Maruz kalma		
	Var	Yok	Toplam
Grup 1	n_{11}	n_{12}	n_{1+}
Grup 2	n_{21}	n_{22}	n_{2+}
Toplam	n_{+1}	n_{+2}	n

$$\tilde{n}_{ij} = n_{ij} + 0.5 \quad i,j = 1,2 \tag{3}$$

eşitliği kullanılarak elde edilen yeni n_{ij} değerlerinin kullanılması önerilmektedir.[14] Agresti tarafından önerilen bağımsız-düzeltilmiş logit güven aralığının hesaplanmasında ise $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarının yerine,

$$\tilde{n}_{ij} = n_{ij} + 2 * (n_{i+}n_{+j}/n^2) \tag{4}$$

eşitliği kullanılarak elde edilen yeni n_{ij} değerlerinin kullanılması önerilmektedir.[15] Gerek Gart gerekse Agresti tarafından önerilen yöntemler dikkate alındığında, Eşitlik 2 ile verilen güven aralığı $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarından bir veya birkaçı 0 değerine eşit olsa bile hesaplanabilir. Ancak $n_{11} = 0$ ve $n_{21} \neq 0$ ise güven aralığının alt sınırı 0'dan büyük olmasına rağmen OR'nin değeri 0 olur. $n_{11} = n_{1+}$ ve $n_{21} = n_{2+}$ veya $n_{11} = 0$ ve $n_{21} = 0$ ise güven aralığının üst sınırı hesaplanabilir olmasına rağmen OR'nin değeri sonsuz olur.[16] Yüzde 95 güven düzeyinde, güven aralığının alt ve üst değerleri için Eşitlik 5 ile verilen formül kullanılır.

$$Alt\ Sınır = e^{\left[\ln(\widehat{OR}) - 1.96 * \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}\right]} \tag{5}$$

$$Üst\ Sınır = e^{\left[\ln(\widehat{OR}) + 1.96 * \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}\right]}$$

Eşitlik 5'te yer alan "e" doğal log için matematiksel sabitken, "ln" doğal log'dur. $n_{11}, n_{12}, n_{21}\ veya\ n_{22}$ frekanslarından herhangi birinin değeri 0 ise OR hesaplanamaz. Bu sorunu düzeltmek için birçok yazar, 2×2 tablonun 0 değeri içermesi durumunda $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarına 0,5 eklenmesini önermektedir.[8,12] Sonuçların sunumunda güven aralıklarının nokta tahminlerine göre önemli bir avantajı, etkinin büyüklüğü ve kesinliği hakkında bilgiyi aynı anda iletmesidir. Güven aralığı basitçe nokta tahmini hakkında hata payı olarak yorumlanır. Güven aralığı ile ilgili önerilen yöntemlerin değerlendirilmesinde kullanılan kriterler:

- Güven aralığının istatistiksel anlamlılığı,
- Güven aralığının hesaplanan OR değerlerini kapsaması,
- Güven aralığının simetrikliği,
- Güven aralığının genişliği,
- Alt ve üst sınır ile ilgili sapmalardan kaçınma,
- Kullanım kolaylığı.

şeklinde sıralanmaktadır.[17] Güven aralığı, OR'nin doğal logaritmasının varyansını tahmin etmeye yönelik bir hesaplamaya dayandığı için yaklaşıktır. Bu yaklaşımın nispeten büyük örneklem büyüklükleri için kesin değerlere oldukça yakın olan güven sınırları verdiği, gerçek OR'nin ve örneklem büyüklüğünün çok büyük olmadığı durumlarda OR için logit güven aralıklarının performansının iyi olduğu ifade edilmektedir.[5,18] Belirli bir maruziyet ile ilgili verilen bir sonuç için pozitif bir OR'nin varlığı, mutlaka bu ilişkinin istatistiksel olarak anlamlı olduğunu göstermez. Önemi belirlemek için güven aralıkları dikkate alınmalıdır.[19]

BULGULAR

Farklı örneklem büyüklükleri için Woolf logit, Gart düzeltilmiş logit ve Agresti bağımsız-düzleştirilmiş logit %95 güven aralıkları Eşitlik 5 kullanılarak belirlenmiş ve Tablo 3A, 3B, 3C'de verilmiştir.

Tablo 3A, 3B, 3C'den de görüldüğü üzere:

TABLO 3A: Woolf logit yöntemi ile ilgili istatistikler.

n	OR	AS	ÜS	$AS < 1 < ÜS$	$AS < OR < ÜS$
10	1,21±2,73	0,12±0,16	49,97±81,28	%100,0	%100,0
15	2,17±5,42	0,23±0,36	53,38±124,95	%90,1	%100,0
20	3,09±8,59	0,35±0,61	58,25±174,03	%82,0	%100,0
25	4,04±11,49	0,50±0,91	61,81±204,43	%72,7	%100,0
30	4,06±14,53	0,53±1,11	57,57±263,85	%67,5	%100,0
35	4,75±15,65	0,68±1,40	55,81±246,06	%62,0	%100,0
40	6,83±25,47	0,91±2,10	83,65±410,96	%59,3	%100,0
45	6,59±25,94	0,95±2,20	76,15±411,69	%57,8	%100,0
50	6,72±25,36	1,05±2,44	67,92±357,78	%51,1	%100,0
55	7,59±22,83	1,24±2,73	65,11±237,99	%53,0	%100,0
60	8,21±41,96	1,28±3,44	84,62±674,56	%50,2	%100,0
65	6,97±34,70	1,23±3,27	62,18±525,18	%46,2	%100,0
70	7,67±29,57	1,42±3,52	59,10±315,50	%44,6	%100,0
75	10,58±44,10	1,83±4,50	88,41±490,44	%43,9	%100,0
80	8,89±43,20	1,60±4,63	71,44±475,40	%43,3	%100,0
85	7,07±23,33	1,54±3,62	45,79±186,12	%41,3	%100,0
90	14,40±88,76	2,27±7,42	145,11±1362,36	%40,1	%100,0
95	10,29±42,65	2,03±5,44	72,95±419,66	%36,4	%100,0
100	9,83±43,19	2,01±5,50	69,13±431,33	%37,7	%100,0
125	16,88±121,76	2,92±11,10	152,86±1709,03	%35,4	%100,0
150	13,10±78,74	2,83±10,39	84,56±699,80	%30,1	%100,0
175	15,16±87,02	3,73±14,74	79,69±588,96	%30,2	%100,0
200	15,18±75,47	3,86±12,33	80,16±588,96	%30,5	%100,0
225	13,90±106,29	3,50±14,02	79,10±933,50	%24,1	%100,0
250	29,90±356,06	5,23±27,89	294,03±5553,94	%23,2	%100,0
300	22,32±232,48	5,15±27,26	142,45±2224,22	%19,3	%100,0
350	13,29±106,88	4,23±15,88	62,37±880,27	%17,9	%100,0
400	25,99±412,37	5,90±48,20	173,84±3885,40	%18,8	%100,0
450	41,44±337,03	9,41±48,30	251,12±2662,44	%18,3	%100,0
500	24,32±350,28	6,67±45,04	139,15±3019,49	%20,0	%100,0
600	23,66±166,98	8,33±36,01	94,45±1249,49	%15,6	%100,0
700	15,81±85,29	6,85±29,15	43,86±284,53	%15,8	%100,0
800	25,12±226,25	8,46±39,84	108,11±1577,13	%13,6	%100,0
900	20,60±152,85	8,43±47,93	59,98±549,87	%15,2	%100,0
1.000	11,27±44,34	5,74±16,92	28,37±221,91	%11,4	%100,0
Genel	13,91±155,74	3,63±23,07	95,77±1633,51	%40,5	%100,0

OR: Odds oranı; AS: Alt sınır; ÜS: Üst sınır, $aritmetik\ ortalama \pm standart\ sapma$.

TABLO 3B: Gart düzeltilmiş logit yöntemi ile ilgili istatistikler.

n	OR	AS	ÜS	$AS < 1 < ÜS$	$AS < OR < ÜS$
10	6,65±17,70	0,18±0,33	313,96±1079,95	%92,3	%78,3
15	8,53±27,88	0,31±0,61	330,03±1544,55	%83,3	%54,1
20	9,13±33,17	0,42±0,86	298,71±1716,35	%75,3	%42,0
25	11,76±51,93	0,58±1,28	389,53±2711,44	%70,2	%32,0
30	11,57±63,63	0,63±1,60	370,25±3256,21	%65,4	%24,6
35	14,90±81,20	0,82±2,07	496,10±4156,01	%62,0	%20,7
40	10,91±45,50	0,92±2,16	209,42±1535,98	%59,2	%15,5
45	14,44±83,84	1,07±2,66	355,45±3992,44	%54,6	%13,7
50	15,41±83,56	1,22±3,03	348,63±3667,91	%49,5	%12,2
55	22,36±171,80	1,47±4,23	714,27±8753,98	%50,2	%12,5
60	20,50±152,10	1,53±4,57	549,22±7263,06	%48,2	%11,6
65	18,02±171,12	1,39±4,35	535,69±8719,74	%45,3	%9,6
70	18,80±189,60	1,55±4,98	556,09±9526,64	%42,6	%8,0
75	24,29±196,70	2,07±6,38	581,61±9158,50	%42,0	%8,3
80	24,16±220,73	1,92±6,49	669,01±10593,48	%42,0	%7,6
85	14,18±67,20	1,70±4,31	199,81±1358,35	%40,9	%8,0
90	21,69±133,49	2,38±7,64	332,66±3117,25	%38,6	%7,5
95	23,11±235,76	2,13±6,50	684,52±12029,43	%35,9	%7,0
100	29,22±323,27	2,45±9,22	782,02±15447,25	%36,7	%6,0
125	39,15±417,05	3,43±13,64	1010,83±19493,55	%35,3	%5,1
150	32,41±396,30	3,20±12,94	863,01±19148,92	%29,0	%4,5
175	23,26±179,18	3,77±14,54	273,10±3454,60	%29,9	%2,7
200	35,20±434,15	4,51±20,59	553,72±10585,22	%29,9	%3,8
225	19,52±129,84	3,61±13,40	199,80±2020,43	%23,5	%3,2
250	30,81±261,48	5,13±24,11	326,91±4135,80	%23,1	%2,4
300	36,10±363,19	5,72±28,95	425,46±6287,55	%19,3	%1,4
350	22,49±159,74	4,57±16,31	232,92±2460,40	%17,7	%1,4
400	28,38±280,44	5,82±40,13	251,77±2865,15	%19,0	%1,6
450	48,95±467,29	9,56±48,07	462,54±7628,14	%18,0	%1,5
500	21,47±218,51	6,23±38,63	124,65±1478,26	%20,3	%2,0
600	38,79±476,99	8,55±37,25	500,61±11623,31	%15,6	%1,2
700	25,59±200,07	7,05±28,78	223,15±3056,09	% 15,9	%1,4
800	24,73±175,48	8,26±37,08	134,90±1546,92	%13,6	%0,7
900	26,47±212,49	8,44±45,00	181,50±2787,43	%15,0	%0,4
1.000	16,30±150,48	5,90±18,45	118,45±2371,44	%11,2	%0,8
Genel	22,96±237,60	3,47±20,87	419,79±7894,32	%39,2	%11,8

OR: Odds oranı; AS: Alt sınır; ÜS: Üst sınır, $aritmetik\ ortalama \pm standart\ sapma$.

TABLO 3C: Agresti bağımsız-düzeltilmiş logit yöntemi ile ilgili istatistikler.

n	OR	AS	ÜS	$AS < 1 < ÜS$	$AS < OR < ÜS$
10	13,72±39,32	0,16±0,33	88300,85±871129,98	%94,3	%100,0
15	17,93±70,92	0,27±0,59	32844399,46±627702754,6	%85,2	%100,0
20	22,20±127,54	0,37±0,82	234782949,9±4696283779	%79,4	%100,0
25	25,74±170,41	0,53±1,24	1,89462E+13±5,82450E+14	%72,9	%100,0
30	23,32±166,76	0,57±1,56	1,49148E+14±3,29467E+15	%68,4	%100,0
35	41,78±420,16	0,75±1,97	2,50163E+19±7,83480E+20	%63,1	%100,0
40	20,84±157,07	0,85±2,11	5,09654E+12±1,59915E+14	%61,1	%100,0
45	23,76±140,05	0,99±2,57	9,99591E+11±3,143218E+13	%58,3	%100,0
50	58,63±1059,16	1,12±2,91	7,36028E+16±2,32065E+18	%52,5	%100,0
55	42,50±364,09	1,38±4,12	8,17764E+21±2,57538E+23	%53,1	%100,0
60	32,11±242,77	1,44±4,46	2,23106E+21±4,96881E+22	%51,2	%100,0
65	32,18±266,78	1,29±4,22	4,02529E+34±1,27100E+36	%48,5	%100,0
70	33,78±356,59	1,48±4,85	8,47215E+22±2,67377E+24	%46,6	%100,0
75	38,94±255,91	1,96±6,25	2,29466E+11±7,20041E+12	%45,2	%100,0
80	55,56±794,99	1,81±6,25	2,48452E+12±7,84884E+13	%44,2	%100,0
85	24,24±131,62	1,60±4,08	7,41176E+14±2,34166E+16	%43,0	%100,0
90	35,65±204,28	2,25±7,58	4,06742E+26±1,28559E+28	%40,7	%100,0
95	56,39±767,38	2,01±5,95	1,01959E+16±3,21705E+17	%38,4	%100,0
100	41,24±393,26	2,36±9,07	5,32248E+12±1,41134E+14	%38,8	%100,0
125	60,34±744,04	3,35±13,39	9,97911E+11±3,15409E+13	%36,7	%100,0
150	79,51±1460,20	3,05±11,96	16444762,69±491967133,2	%31,1	%100,0
175	33,08±274,93	3,70±14,14	6,82804E+12±2,15802E+14	%31,3	%100,0
200	44,85±456,16	4,41±20,37	3,60697E+12±1,13963E+14	%31,5	%100,0
225	34,18±227,56	3,51±13,39	2,17145E+16±6,86328E+17	%25,0	%100,0
250	48,16±409,18	5,09±24,87	1,93117E+16±6,10688E+17	%24,3	% 100,0
300	44,98±400,13	5,65±28,90	115061477,7±3626052760	%20,1	%100,0
350	31,46±246,63	4,48±16,00	47529016866±1,50300E+12	%19,0	%100,0
400	38,79±312,02	5,73±40,39	2846057422±89621311989	%19,7	%100,0
450	63,42±656,14	9,48±47,34	52175153,79±13808554227	%19,0	%100,0
500	31,88±253,94	6,20±39,22	1,47555E+13±4,66609E+14	%20,9	%100,0
600	200,09±5383,36	8,10±33,74	1180667,99±26160177,56	%15,6	%100,0
700	37,59±268,66	6,99±28,66	1,50175E+19±4,74896E+20	%17,0	%100,0
800	36,28±284,83	8,23±37,38	3112400947±98105629647	%14,0	%100,0
900	33,53±277,53	8,43±45,28	88082,23±2583069,20	%15,7	%100,0
1.000	23,82±222,86	5,85±17,98	12655238,14±400006972,8	%11,7	%100,0
Genel	43,24±1039,52	3,40±20,75	1,18650E+33±2,18213E+35	%41,1	%100,0

OR: Odds oranı; AS: Alt sınır; ÜS: Üst sınır, $aritmetik\ ortalama \pm standart\ sapma$.

- Elde edilen güven aralıkları istatistiksel anlamlılık bakımından değerlendirildiğinde; örneklem büyüklüğü arttıkça her üç yöntemden elde edilen güven aralıkları istatistiksel olarak anlamlı hâle gelmektedir. Özellikle $n \geq 150$ olması durumunda, güven aralıklarının istatistiksel anlamlılığı belirginleşmektedir. Genel olarak bakıldığında istatistiksel olarak anlamlı olmayan güven aralığı oranları Woolf logit yöntemi için %40,5, Gart düzeltilmiş logit yöntemi için %39,2 ve Agresti bağımsız-düzleştirilmiş logit yöntemi için ise %41,1'dir.
- Aralıkların OR'lerinin ortalamasına çok yakın veya çok uzak olup olmaması bakımından değerlendirildiğinde; örneklem büyüklüğü arttıkça her üç yöntem için güven aralıkları (özellikle üst sınır değerleri) ortalama odds değerlerinden uzaklaşmaktadır. Genel olarak bakıldığında en az uzaklaşma Woolf logit, en fazla uzaklaşma ise Agresti bağımsız-düzleştirilmiş logit yönteminden elde edilen güven aralıklarında gözlenmiştir.
- Aralık genişliği bakımından değerlendirildiğinde; güven aralıkları hem örneklem büyüklüğü hem de $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekansları ile ilişkilidir. $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarından herhangi birinin değerinin diğerlerine göre aşırı derecede farklılaşması OR'nin değerini önemli derecede etkilemekte, güven aralığını ise büyütmektedir. Örneklem büyüklüğü arttıkça her üç yöntem için güven aralıkları büyümektedir. Genel olarak bakıldığında en dar güven aralıkları Woolf logit yönteminden en geniş güven aralıkları ise Agresti bağımsız-düzleştirilmiş logit yönteminden elde edilmiştir.
- $[0,1]$ dışındaki sınırlar veya 0 genişlik aralıkları gibi sapmalardan kaçınma bakımından değerlendirildiğinde; örneklem büyüklüğü arttıkça her üç yöntem için $[0,1]$ aralıklarının sayısı artmaktadır. Genel olarak bakıldığında $[0,1]$ aralıklarının oranı Woolf logit yöntemi için %27,6, Gart düzeltilmiş logit yöntemi için %31,3 ve Agresti bağımsız-düzleştirilmiş logit yöntemi için ise %30,3'tür. Her üç yöntem için güven aralığının 0'a eşit olması ise söz konusu değildir.
- Kullanım kolaylığı bakımından değerlendirildiğinde, her üç yöntemin de benzer zorlukta olduğu söylenebilir.

TARTIŞMA

İstatistiksel yorum, klinisyen için başlangıçta önemsenmeyebilir. Ancak istatistiklerin dilini anlamak klinisyenin klinik ortamda araştırma bulgularını anlamlı bir şekilde yorumlaması için önemlidir. Güven aralığı ve OR, klinisyene daha iyi bir içgörü sunabilen, tıbbi tedavi ve bakımla ilgili kararları yönlendirmeye yardımcı olabilecek, yaygın olarak kullanılan 2 önemli istatistiktir. Güven aralıkları, birçok istatistiksel analizin doğru yorumlanması için gerekli olan temel bilgileri içermektedir. Bu nedenle güven aralıklarını doğru bir şekilde anlamak ve yorumlamak önemlidir, aksi takdirde yanlış veya yanıltıcı sonuçlara ulaşılması söz konusu olabilir. Güven aralıkları birçok istatistiksel analizin temel çıktısıdır ve parametre tahminlerinin yorumlanmasında kritik bir role sahiptir. Bildirilen aralıktaki tüm değerlerin olası popülasyona ait değer olarak eşit derecede geçerli olduğu anlayışıyla, anlamlı sonuçlar için eşik genellikle %95 olarak alınır.[2] Yaygın olarak kullanılsalar da güven aralıkları genellikle yanlış anlaşılır ve yanlış yorumlanır. Basitçe güven aralığı, tahminin kesin bir tahmin mi yoksa kaba bir tahmin mi olduğunu gösterir. Dar bir güven aralığı kesin bir tahmine işaret ederken, geniş bir güven aralığı daha az kesin bir tahmine işaret eder. Güven aralıkları ve p değerleri farklı bilgiler sağlasa da yakından ilişkilidir. p değerleri hipotez testlerinin sonucu olup, verilerin yokluk hipotezini reddetmek için yeterli kanıt sağlayıp sağlamadığını belirtirken güven aralıkları, gruplar arası farkın ne kadar belirsiz olduğunu açıklar. Güven aralıklarıyla ilgili yaygın bir diğer yanlış anlama, %95'lik bir güven aralığı için gerçek popülasyon ortalamasının alt sınır ile üst sınır arasında yer alma olasılığının %95 olduğudur. Oysa %95 güven aralığı, çalışma aynı popülasyondan gelen çok sayıda örneklem üzerinde yürütülürse, elde edilecek güven aralıklarının %95'inin gerçek popülasyon ortalamasını içermesinin beklenildiği anlamına gelir. Grup karşılaştırmalarında OR için %95 güven aralığı 1 değerini içeriyorsa p değeri 0,05'ten büyük, tersi durumda p değeri kesinlikle 0,05'ten küçük olacaktır.[20] Bilimsel çalışmalarda çoğunluka p değerlerinin kullanılması, araştırmacıları tahmin ve güven aralıkları gibi çalışma sonuçlarını yorumlamaya yönelik daha yararlı yaklaşımlardan uzaklaştırmaktadır. Tıbbi araştırmalarda araştırmacılar, istatistiksel olarak anlamlı olup olmadığına dair basit bir göstergeden ziyade genellikle gruplar arasında ölçülen bir sonucun farkının boyutunu belirlemekle ilgilenirler. Güven aralıkları, böyle bir farkın popülasyon değerinin içinde bulunabileceği, örnek verilere dayalı olarak bir değerler aralığı sunar. Bir güven aralığı, örneklem ortalaması, örneklem ortalamaları arasındaki fark vb. gibi tek bir değer tahmininden popülasyon için makul olduğu düşünülen değer aralığına bir hareket üretir. Örnekleme dayalı güven aralığının genişliği, kısmen standart hataya, örneklem büyüklüğüne ve ortaya çıkan aralıkla ilişkilendirmek istenilen güven derecesine bağlıdır.[21] Kategorik veriler için RR ve OR gibi risk ile ilgili etki büyüklüğü ölçümleri araştırmacılar tarafından sıklıkla kullanılmasına rağmen araştırmacılar tarafından bu istatistiklerin güçlü yönleri ve sınırlamaları tam olarak bilinmemektedir.[22] OR bir ilişki ölçüsü olarak kullanıldığında, birlikteliğin hem gücünü hem de yönünü ifade eder. Yüzde 95 güven aralığı, olası örnekleme hatasının derecesini tahmin eder.[23] OR için güven aralıkları, tipik olarak, OR tahmininin logaritmasının yaklaşık normal dağılımına dayanır. Bu tür aralıklar, logit aralıkları olarak ifade edilir ve bazen gözlemlenen olay ve olay sayılarına yönelik bir ayarlamayı içerir. Güven aralıklarının genel olarak dezavantajı aşırı tutucu olmalarıdır. Özellikle küçük örneklem büyüklüklerinde ve oranlar 0 veya 1'e yakın olduğunda genellikle gerekenden daha geniştirler.[7] OR üstel lojistik regresyon katsayısı olarak tıp, epidemiyoloji ve biyoistatistikte popüler bir ilişki ölçüsüdür. Bir güven aralığı oluştururken, gerçek parametre değerini kapsama olasılığının yanı sıra "yanlış" parametre değerini kapsama olasılığını da hesaba katmak gerekir. Bir bakıma bu değerlendirme, istatistiksel bir testin II. tip hatasının hesaplanmasına benzer.[24] Genel olarak aralık tahmini problemine 2 yaklaşım vardır. Birincisi 2×2 tabloların analizinden oluşmaktadır. İkinci yaklaşım, OR'nin regresyon katsayısı ile doğrudan ilişkili olduğu lojistik modele dayanmaktadır. Ne yazık ki tüm bu güven aralıklarının çok önemli bir dezavantajı, gerçek kapsama olasılıklarının nominal olandan önemli ölçüde daha büyük olmasıdır. Bu durum, Neyman'ın güven aralığı tanımına aykırı olduğundan, yanlış bir sonuca (fazla veya eksik tahmin) ulaşma riski varsayılandan daha fazladır ve ne yazık ki bilinmezliğini korumaktadır.[25]

SONUÇ

Tablo 3A, 3B, 3C'de, çalışmada dikkate alınan örneklem büyüklükleri için Woolf logit, Gart düzeltilmiş logit ve Agresti bağımsız-düzeltilmiş logit yöntemlerine ilişkin bazı istatistikler verilmiştir. Her üç yöntem de gerek OR değeri gerekse güven aralığı bakımından örneklem büyüklüğünden benzer şekilde etkilenmektedir. Örneklem büyüklüğü arttıkça OR değerleri ve güven aralıkları büyümektedir. Ancak hemen tüm örneklem büyüklükleri için Woolf logit yönteminden elde edilen OR değerleri ve güven aralıkları diğer 2 yönteme göre daha küçüktür. $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarından herhangi bir veya birkaçının değerinin diğerlerine göre aşırı derecede farklılaşması yöntemlerden elde edilen OR değerlerinin farklılaşmasına ve özellikle üst sınırların aşırı derecede büyümesine sebep olmaktadır. Bu durumdan Agresti bağımsız-düzeltilmiş logit yöntemi, diğer 2 yönteme göre daha fazla etkilenmektedir. Hemen tüm örneklem büyüklükleri için $alt\ sınır < 1 < üst\ sınır$ durumu her üç yöntemde de genel olarak benzerlik göstermektedir. Genel olarak bakıldığında her üç yöntem için simülasyonların yaklaşık %40'ında $alt\ sınır < 1 < üst\ sınır$ durumu söz konusudur. Hesaplanan OR değerlerinin alt ve üst sınırlar arasında kalması bakımından yöntemler karşılaştırıldığında Woolf logit ve Agresti bağımsız-düzeltilmiş logit yöntemlerinde tüm örneklem büyüklükleri için $alt\ sınır < OR < üst\ sınır$ sağlanmasına rağmen Gart düzeltilmiş logit yöntemi için değişkenlik gösterdiği belirlenmiştir. OR değeri, güven aralığı, $alt\ sınır < 1 < üst\ sınır$ ve $alt\ sınır < OR < üst\ sınır$ kriterleri aynı anda dikkate alındığında çalışmada dikkate alınan yöntemlerden hiçbirinin en iyi yöntem olarak belirtilemeyeceği, bu yöntemlerden yararlanılarak elde edilen güven aralıklarının çok büyük ve nokta tahminleri konusunda simetrik olmadığı tespit edilmiştir. Özellikle $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarında aşırı dengesizliğin söz konusu olduğu durumlarda tahmin edilen OR değerlerinin aşırı derecede büyük olmalarından dolayı güven aralıkları büyümektedir. Her üç yöntem için $n \geq 150$ olması durumunda güven aralıklarının 1 değerini içerme oranı %30'un altına düşmekte ve istatistiksel olarak anlamlı hâle gelmektedir. OR için güven aralığı, kolaylıkları ve doğrulukları bakımından farklılık gösteren çeşitli yöntemler kullanılarak oluşturulabilir. Çalışmada dikkate alınan yöntemler birlikte değerlendirildiğinde Woolf yöntemi, hesaplanan OR'nin istatistiksel anlamlılığı, güven aralığının büyüklüğü ve $alt\ sınır < OR < üst\ sınır$ kriterleri dikkate alındığında diğer 2 yönteme göre daha iyi performans gösterdiği için tatmin edici bir yaklaşım olarak tavsiye edilmektedir. Ancak $n_{11}, n_{12}, n_{21}\ ve\ n_{22}$ frekanslarından herhangi birinin 0 olması durumunda Woolf yönteminden elde edilecek güven aralığının anlamsız hâle geldiği göz ardı edilmemelidir.

Finansal Kaynak

Bu çalışma sırasında, yapılan araştırma konusu ile ilgili doğrudan bağlantısı bulunan herhangi bir ilaç firmasından, tıbbi alet, gereç ve malzeme sağlayan ve/veya üreten bir firma veya herhangi bir ticari firmadan, çalışmanın değerlendirme sürecinde, çalışma ile ilgili verilecek kararı olumsuz etkileyebilecek maddi ve/veya manevi herhangi bir destek alınmamıştır.

Çıkar Çatışması

Bu çalışma ile ilgili olarak yazarların ve/veya aile bireylerinin çıkar çatışması potansiyeli olabilecek bilimsel ve tıbbi komite üyeliği veya üyeleri ile ilişkisi, danışmanlık, bilirkişilik, herhangi bir firmada çalışma durumu, hissedarlık ve benzer durumları yoktur.

Yazar Katkıları

Bu çalışma hazırlanırken tüm yazarlar eşit katkı sağlamıştır.

KAYNAKLAR

1. Wang W, Shan G. Exact confidence intervals for the relative risk and the odds ratio. Biometrics. 2015;71(4):985-95. [Crossref] [PubMed] [PMC]
2. George A, Stead TS, Ganti L. What's the risk: differentiating risk ratios, odds ratios, and hazard ratios? Cureus. 2020;12(8):e10047. [Crossref] [PubMed] [PMC]
3. Grant RL. Converting an odds ratio to a range of plausible relative risks for better communication of research findings. BMJ. 2014;348:f7450. Erratum in: BMJ. 2014;348:g2124. [Crossref] [PubMed]
4. Bland JM, Altman DG. Statistics notes. The odds ratio. BMJ. 2000;320(7247):1468. [Crossref] [PubMed] [PMC]
5. Andres AM, Tejedor H, Hernandez MA. Two-tailed asymptotic inferences for the odds ratio in prospective and retrospective studies: evaluation of methods of inference. J Stat Comput Simul. 2020;90(1):138-56. [Crossref]
6. O'Brien SF, Yi QL. How do I interpret a confidence interval? Transfusion. 2016;56(7):1680-3. [Crossref] [PubMed]
7. Fagerland MW. Exact and mid-p confidence intervals for the odds ratio. Stata J. 2012;12(3):505-14. [Crossref]
8. Tenny S, Hoffman MR. Odds Ratio. 2022. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022. [PubMed]
9. Yıldırım HH, Yıldırım S. Hipotez testi, güven aralığı, etki büyüklüğü ve merkezi olmayan olasılık dağılımları üzerine [On hypothesis testing, confidence interval, effect size and noncentral probability distributions]. İlköğretim Online. 2011;10(3):1112-23. [Link]
10. Laing CM, Rankin JA. Odds ratios and confidence intervals: a review for the pediatric oncology clinician. J Pediatr Oncol Nurs. 2011;28(6):363-7. [Crossref] [PubMed]
11. Simon SD. Understanding the odds ratio and the relative risk. J Androl. 2001;22(4):533-6. [PubMed]
12. Islam N. Symmetry of odds ratio. Int J Phys Soc Sci. 2013;3(12):580-2. [Link]
13. Woolf B. On estimating the relation between blood group and disease. Ann Hum Genet. 1955;19(4):251-3. [Crossref] [PubMed]
14. Gart JJ. Alternative analyses of contingency tables. JR Stat Soc Series B Methodol. 1966;28(1):164-79. [Crossref]
15. Agresti A. On logit confidence intervals for the odds ratio with small samples. Biometrics. 1999;55(2):597-602. [Crossref] [PubMed]
16. Fagerland MW, Lydersen S, Laake P. Recommended confidence intervals for two independent binomial proportions. Stat Methods Med Res. 2015;24(2):224-54. [Crossref] [PubMed]
17. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med. 1998;17(8):857-72. [Crossref] [PubMed]
18. Sadinle M. Transformed logit confidence intervals for small populations in single capture-recapture estimation. Commun Stat Simul Comput. 2009;38(9):1909-24. [Crossref]
19. Szumilas M. Explaining odds ratios. J Can Acad Child Adolesc Psychiatry. 2010;19(3):227-9. Erratum in: J Can Acad Child Adolesc Psychiatry. 2015;24(1):58. [PubMed] [PMC]
20. Tan SH, Tan SB. The correct interpretation of confidence intervals. Proc Singap Healthc. 2010;19(3):276-8. [Crossref]
21. Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. Br Med J (Clin Res Ed). 1986;292(6522):746-50. [Crossref] [PubMed] [PMC]
22. Andrade C. Understanding relative risk, odds ratio, and related terms: as simple as it can get. J Clin Psychiatry. 2015;76(7):e857-61. [Crossref] [PubMed]
23. Sandercock P. The odds ratio: a useful tool in neurosciences. J Neurol Neurosurg Psychiatry. 1989;52(7):817-20. [Crossref] [PubMed] [PMC]
24. Demidenko E. The shortest width confidence interval for odds ratio in logistic regression. Open J Stat. 2012;2(3):305-8. [Crossref]
25. Zielinska-Kolasinska Z, Zielinski W. A new confidence interval for odds ratio. arXiv:1910.03832v2[stat.ME]. 2020. [Crossref]