

ORJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2025-111374

Tanısal Olasılık Oranı, Etki Eden Faktörler ve Diğer Tanı Performans Kriterleri Arasındaki İlişkinin İncelenmesi: Bir Simülasyon Çalışması

Examining the Relationship Between Diagnostic Odds Ratio, Influencing Factors and Other Diagnostic Performance Criteria: A Simulation Study

İsmet DOĞAN^a, Nurhan DOĞAN^a

^aAfyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik AD, Afyonkarahisar, Türkiye

ÖZET Amaç: Bu çalışmanın amacı, örneklem büyüklüğünün tanısal olasılık oranı [diagnostic odds ratio (DOR)] değerleri ve güven aralıkları üzerindeki etkisini belirlemek, DOR ile örneklem büyüklüğü ve diğer tanı performans göstergeleri arasındaki ilişkileri ortaya koymaktır. **Gereç ve Yöntemler:** Tanısal test performans kriterlerinden biri olan DOR hesaplamaları için farklı örneklem büyüklükleri dikkate alınarak bir simülasyon çalışması yapılmıştır. Python-random kütüphanesi kullanılarak $10 \leq n \leq 1.000$ aralığında 35 farklı n değeri için veri elde edilmiş, 34.101 farklı 2×2 çapraz tablo değerlendirilmiştir. Farklı örneklem büyüklüklerine göre elde edilen DOR ve etkileyen faktörler ile diğer performans kriterleri arasındaki ilişkilerin değerlendirilmesinde Spearman korelasyon katsayısı kullanılmıştır. **Bulgular:** DOR değerleri ve güven aralıkları örneklem büyüklüğünden etkilanmemektedir. Örneklem büyüklüğü ile DOR arasında bir ilişki olup olmadığı incelenmiş ve genel olarak pozitif ancak düşük düzeyde $r=0,063$ bir ilişki tespit edilmiştir. $DOR < 1$ ve $DOR > 1$ olması neredeyse %50-50 oranında birbirine eşit olarak belirlenmiştir. DOR değerleri yalnızca negatif ve yalnızca pozitif değerleri ile negatif, gerçek pozitif ve gerçek negatif değerleri ile pozitif ilişkiye sahiptir. DOR ile negatif olabilirlik oranı arasında negatif diğer metrikler ile pozitif ilişki olduğu belirlenmiştir. **Sonuç:** DOR, tıbbi tanı testlerinin doğruluğunu ölçmede etkili ve özetleyici bir göstergedir. Testin performansını tek bir değerle ifade edebilmesi, özellikle tanı testlerinin karşılaştırılması ve metaanaliz çalışmalarında önemli bir referans noktası olmasını sağlamaktadır. DOR'un değerinin büyümesi, testin hem hastaları doğru şekilde tespit etme hem de sağlıklı bireyleri doğru şekilde dışlama konusunda başarılı olduğunu gösterir. Ancak, daha güvenilir sonuçlar elde edebilmek için tek başına değerlendirilmemeli, duyarlılık ve özgüllük gibi diğer metriklerle birlikte ele alınmalıdır.

ABSTRACT Objective: The aim of this study is to determine the effect of sample size on diagnostic odds ratio (DOR) values and confidence intervals, and to determine the relationships between DOR and sample size and other diagnostic performance indicators. **Material and Methods:** A simulation study was performed considering different sample sizes for the calculation of the DOR, which is one of the performance criteria for diagnostic tests. Using the Python-random library, data were obtained for 35 different n values in the range $10 \leq n \leq 1,000$, and 34,101 different 2×2 cross tables were evaluated. The Spearman correlation coefficient was used to evaluate the relationships between the DOR obtained according to different sample sizes and the influencing factors and other performance criteria. **Results:** Neither the DOR values nor the confidence intervals are influenced by sample size. An analysis was conducted to assess the relationship between sample size and DOR, revealing a generally positive, though weak $r=0.063$, correlation. The occurrence of $DOR < 1$ and $DOR > 1$ was found to be nearly equal, with a distribution of 50-50%. DOR values exhibited a negative correlation with false negatives and false positives, while showing a positive correlation with true positives and true negatives. Additionally, a negative correlation between DOR and the negative likelihood ratio rate was observed, while a positive correlation with other performance metrics was identified. **Conclusion:** DOR is an effective and summative indicator for measuring the accuracy of medical diagnostic tests. The fact that the test can express its performance with a single value makes it an important reference point, especially in comparisons of diagnostic tests and metaanalysis studies. The increase in the value of DOR indicates that the test is successful in both correctly identifying patients and correctly excluding healthy individuals. However, to ensure more reliable results, DOR should not be evaluated in isolation, but rather in conjunction with other metrics such as sensitivity and specificity.

Anahtar kelimeler: Tanı testi; test performansı; tanı ve prognoz çalışmaları; kanıta dayalı tıp

Keywords: Diagnostic test; test performance; diagnostic and prognostic studies; evidence-based medicine

KAYNAK GÖSTERMEK İÇİN:

Doğan İ, Doğan N. Tanısal olasılık oranı, etki eden faktörler ve diğer tanı performans kriterleri arasındaki ilişkinin incelenmesi: Bir simülasyon çalışması. Türkiye Klinikleri J Biostat. 2025;17(1):40-50.

Correspondence: İsmet DOĞAN

Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik AD, Afyonkarahisar, Türkiye

E-mail: ismet.dogan@afsu.edu.tr

Peer review under responsibility of Türkiye Klinikleri Journal of Biostatistics.

Received: 18 Apr 2025

Received in revised form: 13 Jun 2025

Accepted: 19 Jun 2025

Available online: 25 Jun 2025

2146-8877 / Copyright © 2025 by Türkiye Klinikleri. This is an open

access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Tanı testleri, sağlık hizmetlerinde hayati bir rol oynar ve doktorların hastanın sağlığı hakkında güvenilir bilgi sağlamasına ve bir tedavi planı önermesine olanak tanır. İyi bir tanı testi, hasta ve sağlıklı bireyler arasında ayırım yapmak için gereklidir.¹ Sağlık profesyonelleri, belirsizlik içeren durumlarda karar vermek zorundadır. Bir hastaya teşhis koyarken, hastanın yaşadığı durumu belirlemek için ilgili özellikler belirli kategorilere ayrılır. Bu süreçte, anamnez, fiziksel muayene ve ek testler gibi yöntemler, bu kategorileri destekleyici rol oynar. Tanı testleri olarak adlandırılan bu araçlar, sağlık uzmanlarının şüphelenilen hastalığın var olup olmadığını değerlendirmesine yardımcı olur. Ancak, tanı testlerinin etkinliği her klinik duruma göre değişiklik gösterebilir. Bu nedenle, doğruluk (duyarlılık ve özgüllük) ile tanı etkisini (sağlık sonuçlarına olan etkisi) değerlendiren çalışmalar yapılmaktadır.² Kanıta dayalı tıbbın ön planda olduğu günümüzde karar vericilerin, bireylerin hastalığa sahip olup olmadığını doğru bir şekilde belirleyebilmesi için güvenilir test sonuçlarına ihtiyacı vardır. Testlerin performansını değerlendirmek amacıyla duyarlılık, özgüllük, öngörü değerleri ve alıcı işletim karakteristiği [receiver operating characteristic (ROC)] eğrisinin altındaki alan gibi çeşitli göstergeler geliştirilmiştir. Bu göstergeler arasında tanısal olasılık oranı [Diagnostic Odds Ratio (DOR)], tıbbi testlerin performansını tek bir değerle özetleyen önemli bir metriktir. DOR, duyarlılık ve özgüllükle yakından bağlantılı olup, gerçek pozitifler ile gerçek negatifler arasındaki ilişkinin gücünü ölçer. Ancak, duyarlılık ve özgüllük birbirine yakın değerlere sahip olduğunda, DOR'nun etkinliği azalabilir.³ DOR, duyarlılık, özgüllük, pozitif ve negatif öngörü değerleri, olabilirlik oranları ve ROC eğrisi altındaki alan [area under the curve (AUC)] gibi tanısal test performansını değerlendiren birçok spesifik nicel göstergeyle birlikte klinik uygulamalarda kullanılmaktadır.⁴ Test performansını ölçen bir metrik olarak DOR, duyarlılık ve özgüllüğün güçlü yönlerini doğruluk avantajıyla tek bir gösterge hâlinde birleştirir. Hastalığı olan bireylerde testin pozitif çıkma olasılığının, hastalığı olmayan bireylerde pozitif çıkma olasılığına oranını ifade eder. Değeri 0 ile sonsuz arasında değişebilir ve daha yüksek değerler, ayırt etme konusunda daha iyi bir performansa sahip olduğunu gösterir. Örneğin, DOR=36, hastalığı bulunan bireylerin test sonucunun pozitif çıkma olasılığının, hastalığı bulunmayanlara kıyasla 36 kat daha yüksek olduğunu ifade eder. DOR=1 olması testin hastalığı olan ve olmayan hastalar arasında ayırım yapmadığı anlamına gelir. DOR<1 olması durumunda ise test uygun değildir. Bu, hastalıklı popülasyonda daha fazla negatif test anlamına gelir. 2×2 tablo sıfır içeriyorsa, DOR tanımsızdır. Bu durumda tablodaki tüm sayılara 0,5 eklemek, DOR'un yaklaşık bir değerini hesaplamak için yaygın olarak kullanılan bir yöntemdir.^{5,6} DOR'un avantaj ve dezavantajları;

- Duyarlılık ve özgüllüğü birleştirerek tek bir performans ölçüsü sağlar,
- Hesaplanması ve yorumlanması kolaydır,
- Farklı tanı testlerinin performansını karşılaştırmak için kullanılabilir,
- Genel test doğruluğunun istatistiksel bir ölçümüdür,
- 2×2 tablosundan kolayca elde edilebilir,
- Cut-off değerinden sıklıkla etkilenmez,
- Hastalığın mutlak riski hakkında bilgi sağlamaz,
- Klinik uygulamada uygulanması kolay değildir,
- Test edilen popülasyondaki hastalığın yaygınlığından etkilenebilir,
- Yüksek duyarlılığa sahip ancak düşük özgüllüğe sahip (veya tam tersi) testler için uygun olmayabilir,
- Klinik kohortların kullanıldığı çalışmalara kıyasla vaka-kontrol çalışmalarında önemli ölçüde abartılıdır,
- Farklı referans testlerin kullanılmasından etkilenir,
- Geriye dönük veri toplama ve hastaların seçilerek çalışmaya dâhil edilmesinden etkilenmez,

şeklinde ifade edilmektedir.^{7,8} Bu çalışmada, yalnızca 2 olası sonuca (pozitif ve negatif) sahip en basit 2'li test türü ele alınmıştır. Bu çalışmanın amacı, örneklem büyüklüğünün DOR değerleri ve güven aralıkları üzerindeki etkisini belirlemek, DOR ile örneklem büyüklüğü ve diğer tanı performans göstergeleri arasındaki ilişkileri ortaya koymaktır. Çalışmada, Helsinki Deklarasyonu prensipleri dikkate alınmıştır.

GEREÇ ve YÖNTEMLER

Tanısal araştırmalarda, örneklem büyüklüğü hesaplaması geçerlilik ve güvenilirliği sağlamada önemli bir rol oynar. Bir tanı testinin doğruluğunu tahmin etmede, doğruluğun önceden belirlenmiş bir hassasiyete göre tahmin edildiğinden emin olmak için örneklem büyüklüğü hesaplanmalıdır. Tanısal doğruluk çalışmalarında doğru örneklem büyüklüğü, çalışma amacına göre duyarlılık, özgüllük, ROC eğrisi vb. doğruluk metriklerine dayanarak hesaplanır.⁹ Çalışmada, Python-random kütüphanesi (Python Software Foundation, USA) 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 simülasyon kurgusunun detayları aşağıdaki gibidir:

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

Adım 2) Python random kütüphanesindeki choice fonksiyonu kullanılarak rastgele bir şekilde kategori seçimi (GP: Gerçek pozitif, GN: Gerçek negatif, YP: Yalancı pozitif, YN: Yalancı negatif) 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'ncü kategori için bir tam sayı seçildikten sonra seçilen sayıların toplamına bakılır. Eğer bu toplam, istenilen toplamdan fazla ise bu geçersiz bir veri olduğundan 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 birinci adımdan itibaren program bir döngü içinde tekrar çalıştırılır. (Bu çalışmada, standardizasyonu sağlamak amacıyla her bir örneklem büyüklüğü için 1.000 adet veri türetilmesi planlanmıştır. Ancak, toplamı 10-15 olan ve tekrar etmeyen 1.000 tane veri bulmak imkânsız olduğundan n=10 için 286, n=15 için 815 adet olmak üzere toplam 34.101 farklı veri hesaplamalarda kullanılmıştır.)

Adım 7) Son olarak türetilen verilerden basit Excel (Microsoft Corporation, USA) uygulaması ile DOR ve diğer performans kriterlerine ait değerler hesaplanır.

Tanısal bir test ile hastalığın varlığı arasındaki ilişki, geleneksel olarak 4 olası test sonucunu içeren ve [Tablo 1](#)'de verilen 2×2 'lik bir olasılık tablosu ile tanımlanır.

Klinisyenler çeşitli klinik ortamlarda tanı verilerinin doğru yorumlanmasına güvenmelidir. 2×2 'lik bir tablo, tanı testi değerlendirmesi için epidemiyolojik çalışmalara ilişkin verileri sunmak için olmazsa olmaz bir araçtır. [Tablo 1](#)'de verilen bir 2×2 tablodan türetilen test performansının yaygın göstergeleri [Tablo 2](#)'de verilmiştir.

TABLO 1: Test performans ölçümlerinin hesaplanmasında kullanılan 2×2 tablosu

		Gerçek hastalık durumu	
Medikal test sonucu		Hasta	Sağlıklı
	Pozitif	GP	YP
	Negatif	YN	GN

GP: Gerçek pozitif; YN: Yalancı negatif; YP: Yalancı pozitif; GN: Gerçek negatif

TABLO 2: Tanı araştırmalarında yaygın olarak kullanılan test göstergeleri

Test göstergesi	Formül	Açıklama
Duyarlılık	$GP/(GP + YN)$	Hastalar arasında pozitif test sonuçlarının oranı
Özgüllük	$GN/(GN + YP)$	Sağlıklı kişiler arasında negatif test sonuçlarının oranı
Pozitif öngörü değeri (PÖD)	$GP/(GP + YP)$	Pozitif test sonucu olan denekler arasında hastalıklıların oranı
Negatif öngörü değeri (NÖD)	$GN/(GN + YN)$	Negatif test sonucu olan denekler arasında hastaliksız olanların oranı
Pozitif olabilirlik oranı (LR+)	$duyarlılık/(1 - özgüllük)$	Hastalarda pozitif test sonucunun sağlıklı kişilerde aynı sonuca oranı
Negatif olabilirlik oranı (LR-)	$(1 - duyarlılık)/özgüllük$	Hastalarda negatif test sonucunun sağlıklı kişilerde aynı sonuca oranı
Doğruluk	$(GP + GN)/(GP + GN + YP + YN)$	Doğru tanımlanan deneklerin oranı
Youden endeksi	$duyarlılık + özgüllük - 1$	

GP: Gerçek Pozitif, YN: Yalancı Negatif, GN: Gerçek Negatif, YP: Yalancı Pozitif, PÖD: Pozitif öngörü değeri, NÖD: Negatif öngörü değeri, LR: Olasılık oranı (likelihood ratio)

Bir tanı testinin doğruluğu klinik uygulamada kullanılmadan önce değerlendirilmelidir. Tanı testlerinin doğruluğunun çeşitli ölçütleri vardır ve karşılaştırma yönteminin seçimi test sonuçlarının doğasına bağlıdır. Sonuçlar 2'li veya sıralı ölçek verileri gibi nitel veya sürekli ölçek verileri gibi nicel olarak sınıflandırılabilir.⁹ DOR, hastalığı olan bireylerde testin pozitif çıkma olasılığının, hastalığı olmayan bireylerde testin pozitif çıkma olasılığına oranıdır. Bu, eşitlik 1-3 ile verilen belirli istatistiksel ilişkilerin geçerli olduğu anlamına gelir.^{3,5}

$$DOR = \frac{GP}{YN} \bigg/ \frac{YP}{GN} = \frac{duyarlılık}{(1 - duyarlılık)} \bigg/ \frac{(1 - özgüllük)}{özgüllük} \dots \dots \dots (1)$$

Alternatif olarak DOR, test pozitiflerindeki hastalık olasılığının test negatiflerindeki hastalık olasılığına oranı olarak da okunabilir.

$$DOR = \frac{GP}{YN} \bigg/ \frac{YP}{GN} = \frac{PPV}{(1 - PPV)} \bigg/ \frac{(1 - NPV)}{NPV} \dots \dots \dots (2)$$

DOR ile olabilirlik oranları arasında da yakın bir ilişki vardır:

$$DOR = \frac{GP}{YN} \bigg/ \frac{YP}{GN} = \frac{LR(+)}{LR(-)} \dots \dots \dots (3)$$

Eşitliklerden de anlaşılacağı üzere, DOR testin kullanıldığı hastalığın yaygınlığına bağlı değildir. Bununla birlikte, klinik uygulamalarda, test performansının diğer tüm göstergelerinde olduğu gibi, hastalığın şiddetinin çeşitliliğine bağlı olması muhtemeldir.⁵ Tanı testleriyle yaygın olarak kullanılan terimler, testin performansını istatistiksel olarak ölçen duyarlılık, özgüllük ve doğruluktur. DOR, bir tanı testinin ayırt edici gücünü ölçmek için kullanılır. DOR, yaygınlığa bağlı değildir ve bilindik bir epidemiyolojik ölçü olduğu için anlaşılması daha kolay olabilir.⁴ Güven aralığı tahminleri Eşitlik 4-5 kullanılarak hesaplanabilir.

$$SE(\ln DOR) = \sqrt{\frac{1}{GP} + \frac{1}{GN} + \frac{1}{YP} + \frac{1}{YN}} \dots \dots \dots (4)$$

log DOR'un %95'lik güven aralığı daha sonra şu şekilde elde edilebilir:

$$\ln DOR \mp 1,96 * SE(\ln DOR) \dots \dots \dots (5)$$

Eşitlik 5'ten elde edilen alt ve üst sınır değerleri için e^x fonksiyonu kullanılarak DOR'un güven aralığı elde edilir.⁵ DOR değerlerinin normal dağılıma sahip olup olmadıkları Kolmogorov-Smirnov testi ile değerlendirilmiş, verilerin normal dağılıma sahip olmadıkları belirlenmiştir ($p>0,05$). Dolayısıyla, farklı örneklem büyüklüklerine göre elde edilen DOR ve etkileyen faktörler ile diğer performans kriterleri arasındaki ilişki-

rin değerlendirilmesinde Spearman korelasyon katsayısı kullanılmıştır. Çapraz tablolarda gözelerin herhangi birinde sıfır değerinin olması durumunda DOR değerlerinin tanımsız olma durumunun giderilmesi için sıfır değerli gözeler 0,5 eklemesi yapılmıştır. Örneklem büyüklüklerine göre düzeltme yapılan 2×2 çapraz tablo sayıları ve yüzdeleri [Tablo 3](#)'te verilmiştir.

[Tablo 3](#)'te de görüldüğü gibi örneklem büyüklüğü arttıkça sıfır değeri içeren çapraz tablo sayısı azalmaktadır. DOR değerlerinin tahmin edilmesinde basit rastgele örnekleme temeline dayalı yöntem esas alınmış olup hesaplamalarda $DOR = \frac{GP}{YN} / \frac{YP}{GN}$ eşitliği kullanılmıştır.

TABLO 3: Örneklem büyüklüklerine göre 0,5 düzeltmesi yapılan 2×2 tablo sayıları

n	Düzeltilme yapılan tablo sayısı	Düzeltilme yapılmayan tablo sayısı	Toplam tablo sayısı
10	202 (%70,6)	84 (%29,4)	286
15	451 (%55,3)	364 (%44,7)	815
20	451 (%45,1)	549 (%54,9)	1.000
25	395 (%39,5)	605 (%60,5)	1.000
30	329 (%32,9)	671 (%67,1)	1.000
35	290 (%29,0)	710 (%71,0)	1.000
40	249 (%24,9)	751 (%75,1)	1.000
45	235 (%23,5)	765 (%76,5)	1.000
50	210 (%21,0)	790 (%79,0)	1.000
55	176 (%17,6)	824 (%82,4)	1.000
60	165 (%16,5)	835 (%83,5)	1.000
65	165 (%16,5)	835 (%83,5)	1.000
70	136 (%13,6)	864 (%86,4)	1.000
75	150 (%15,0)	850 (%85,0)	1.000
80	128 (%12,8)	872 (%87,2)	1.000
85	148 (%14,8)	852 (%85,2)	1.000
90	130 (%13,0)	870 (%87,0)	1.000
95	122 (%12,2)	878 (%87,8)	1.000
100	120 (%12,0)	880 (%88,0)	1.000
125	92 (%9,2)	908 (%90,8)	1.000
150	81 (%8,1)	919 (%91,9)	1.000
175	67 (%6,7)	933 (%93,3)	1.000
200	61 (%6,1)	939 (%93,9)	1.000
225	58 (%5,8)	942 (%94,2)	1.000
250	45 (%4,5)	955 (%95,5)	1.000
300	34 (%3,4)	966 (%96,6)	1.000
350	39 (%3,9)	961 (%96,1)	1.000
400	33 (%3,3)	967 (%96,7)	1.000
450	29 (%2,9)	971 (%97,1)	1.000
500	33 (%3,3)	967 (%96,7)	1.000
600	20 (%2,0)	980 (%98,0)	1.000
700	33 (%3,3)	967 (%96,7)	1.000
800	12 (%1,2)	988 (%98,8)	1.000
900	12 (%1,2)	988 (%98,8)	1.000
1.000	11 (%1,1)	989 (%98,9)	1.000
Genel	4.912 (%14,4)	29.189 (%85,6)	34.101

BULGULAR

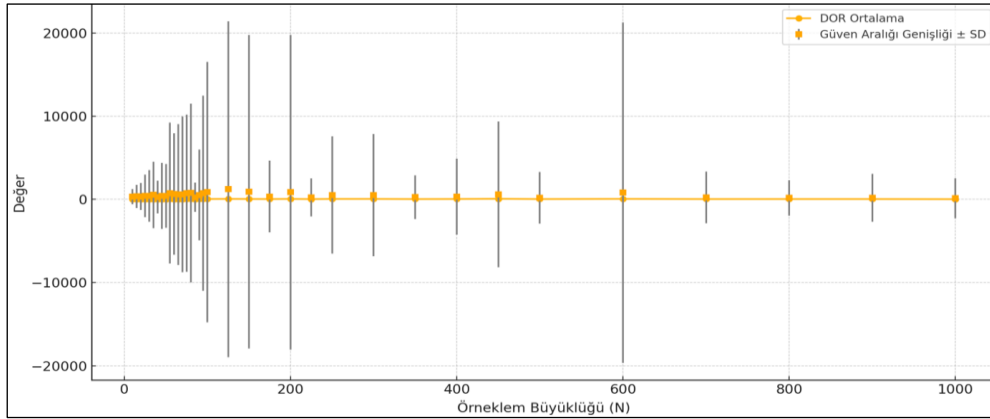
Bir tanı prosedürünün veya testin doğruluğu, “Bu test, belirli 2 durumu (örneğin, sağlık ve hastalık ya da bir hastalığın farklı evreleri) ne kadar iyi ayırt edebilir?” sorusuna yanıt verir. Bu ayırt edici yetenek, tanısal doğruluk ölçütleri aracılığıyla değerlendirilebilir.¹⁰ Farklı örneklem büyüklükleri için çalışmada elde edilen DOR değerlerine ilişkin istatistikler [Tablo 4](#)’te verilmiştir.

[Tablo 4](#)’ten de görüldüğü üzere gerek DOR değerleri gerekse güven aralıkları düzenli bir biçimde örneklem büyüklüğünden etkilenmemektedir. Bu durum, [Şekil 1](#)’de de açıkça görülmektedir.

TABLO 4: Örneklem büyüklüklerine göre DOR ile ilgili istatistikler

n	X±SS	DOR<1	DOR=1	DOR>1	Alt sınır	Üst sınır
10	6,265±14,920	%46,5	%7,0	%46,5	0,152±0,282	312,049±909,973
15	8,789±25,833	%49,1	%2,0	%49,0	0,288±0,581	348,441±1387,378
20	10,011±33,112	%48,2	%2,3	%49,5	0,417±0,883	334,867±1619,398
25	12,818±50,668	%48,6	%1,0	%50,4	0,591±1,320	417,581±2541,910
30	12,867±63,464	%51,2	%1,6	%47,2	0,647±1,685	404,399±3111,618
35	16,392±81,288	%50,5	%0,5	%49,0	0,854±2,192	528,755±3976,918
40	13,899±60,079	%49,2	%1,1	%49,7	0,987±2,525	291,382±1951,209
45	17,034±89,354	%49,0	%0,7	%50,3	1,138±2,948	418,319±3980,894
50	18,378±93,847	%50,1	%0,6	%49,3	1,304±3,408	425,094±3819,568
55	24,661±169,194	%49,4	%0,3	%50,3	1,563±4,410	751,697±8473,378
60	24,355±162,354	%49,7	%0,6	%49,7	1,648±5,107	645,993±7293,077
65	20,105±169,344	%50,2	%0,3	%49,5	1,470±4,603	574,492±8466,848
70	21,206±191,090	%51,2	%0,6	%48,2	1,657±5,325	600,285±9356,037
75	29,407±214,394	%48,8	%0,2	%51,0	2,257±7,158	713,485±9433,505
80	28,399±235,530	%50,8	%0,2	%49,0	2,064±7,197	774,702±10735,682
85	16,398±78,076	%49,0	%0,0	%51,0	1,807±4,782	241,290±1743,540
90	29,172±196,955	%48,7	%0,2	%51,1	2,632±9,199	524,465±5440,133
95	25,871±234,522	%48,8	%0,2	%51,0	2,271±7,001	721,921±11710,487
100	33,739±340,870	%53,3	%0,1	%46,6	2,638±10,070	883,832±15631,200
125	47,997±461,130	%50,5	%0,0	%49,5	3,765±15,579	1229,166±20167,167
150	36,662±404,039	%50,0	%0,0	%50,0	3,457±14,394	919,489±18814,826
175	27,929±216,441	%52,1	%0,1	%47,8	4,145±16,973	338,260±4296,257
200	45,064±636,069	%48,5	%0,2	%51,3	4,894±24,433	853,311±18871,721
225	22,811±157,464	%51,2	%0,1	%48,7	3,837±15,076	236,075±2276,394
250	42,252±417,699	%52,1	%0,0	%47,9	5,691±29,476	524,382±7008,660
300	43,309±436,755	%51,3	%0,0	%48,7	6,165±33,170	521,069±7308,709
350	25,077±181,108	%51,4	%0,0	%48,6	4,783±17,740	259,654±2617,2672
400	35,758±432,757	%49,7	%0,0	%50,3	6,327±48,109	338,534±4522,198
450	61,057±561,124	%47,7	%0,0	%52,3	10,271±53,385	603,779±8695,075
500	27,204±350,093	%49,9	%0,0	%50,1	6,643±44,328	187,848±3050,697
600	49,126±687,562	%47,3	%0,0	%52,7	9,125±41,820	811,792±20424,462
700	27,213±205,746	%49,8	%0,0	%50,2	7,343±30,761	232,064±3078,457
800	29,257±239,992	%47,8	%0,0	%52,2	8,618±39,861	176,903±2069,215
900	28,844±228,386	%50,5	%0,0	%49,5	8,841±48,690	194,636±2818,629
1.000	17,165±152,975	%50,0	%0,0	%50,0	6,047±19,009	124,791±2384,786
Genel	27,283±296,514	%49,8	%0,4	%49,8	3,700±23,441	503,725±9016,852

DOR: Tanısal olasılık oranı (diagnostic odds ratio); SS: Standart sapma



ŞEKİL 1: Örneklem büyüklüklerine göre DOR ve güven aralıkları değişimi

DOR: Tanısal olasılık oranı (diagnostic odds ratio); SD: Standart sapma

[Şekil 1](#) incelendiğinde, küçük örneklem büyüklüklerinde DOR değerlerinin ortalama ve varyans açısından oldukça değişken olduğu söylenebilir. Bu durum, düşük örneklem büyüklükleri için elde edilen DOR tahminlerinde istatistiksel belirsizliğin arttığını göstermektedir. Örneklem büyüklüğü arttıkça DOR ortalamaları daha stabil bir seyir izlemekte ve güven aralığı genişlikleri belirgin şekilde azalmaktadır. Özellikle 300'ün üzerindeki örneklem büyüklüklerinde hata çubuklarının daralması, DOR hesaplamalarının daha güvenilir hâle geldiğini göstermektedir. Çalışmada elde edilen DOR tahmin değerlerinin tamamı %95'lik güven aralıkları içerisinde kalmıştır. $DOR < 1$ ve $DOR > 1$ olması neredeyse %50-50 oranında birbirine eşit olarak belirlenmiştir. Tanısal doğruluğun farklı ölçümleri, tanı sürecinin çeşitli yönlerini ele alır. Bazı ölçümler, testin ayırt edici gücünü değerlendirmeye yönelikken, diğerleri testin öngörü yeteneğini analiz etmeye odaklanır. Ayırt edici ölçümler genellikle sağlık politikası kararlarında kullanılırken, öngörücü ölçümler bireysel hastalık olasılıklarını tahmin etmek için daha faydalıdır. Ayrıca, bir testin performansına ilişkin ölçütlerin sabit göstergeler olmadığı unutulmamalıdır. Tanısal doğruluk ölçümleri, testin uygulandığı popülasyonun özelliklerine oldukça duyarlıdır. Bazıları hastalık prevalansından önemli ölçüde etkilenirken, diğerleri çalışılan popülasyondaki hastalık çeşitliliğine karşı hassastır. Bu nedenle, bu ölçümlerin nasıl yorumlanacağını, hangi durumlarda ve hangi koşullar altında kullanılacağını bilmek büyük önem taşır.¹⁰ Fischer ve ark. tanı testleriyle ilgili makalelerin eleştirel değerlendirmesini geliştirmek amacıyla, duyarlılık, özgüllük, ROC eğrisi, olasılık oranları ve testleri karşılaştırmaya yönelik alternatif bir yöntem olarak DOR da dâhil olmak üzere çeşitli test doğruluk ölçümlerinin raporlanmasını önermektedir.¹¹ Çalışmada, örneklem büyüklüğü ile DOR arasında bir ilişki olup olmadığı incelenmiş ve genel olarak pozitif ancak düşük düzeyde $r=0,063$ bir ilişki tespit edilmiştir. Farklı örneklem büyüklükleri için 2×2 tablo değerleri ve performans kriterleri ile DOR değerleri arasındaki ilişki miktarları da hesaplanmış ve [Tablo 5](#)'te verilmiştir.

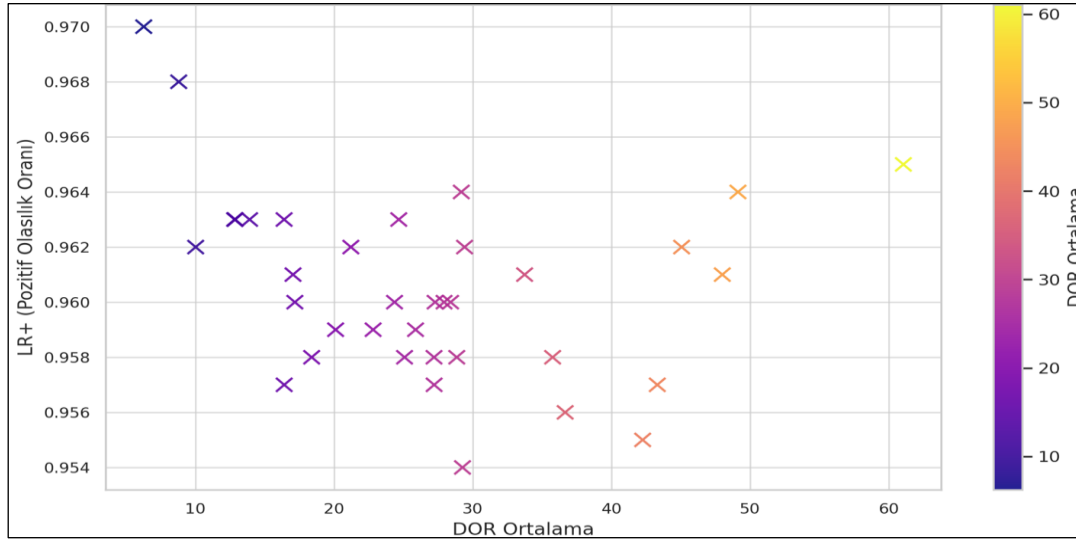
[Tablo 5](#)'te de görüldüğü üzere DOR değerleri yalancı negatif ve yalancı pozitif değerleri ile ters orantılı gerçek pozitif ve gerçek negatif değerleri ile doğru orantılı ilişkiye sahiptir. Performans kriterleri ile DOR arasında ise negatif benzerlik oranı ile ters orantılı diğer kriterler ile doğru orantılı ilişki olduğu belirlenmiştir $p < 0,001$. [Tablo 5](#)'teki korelasyon katsayıları, her bir örneklem büyüklüğü için ayrı ayrı listelenmiştir. Bu, veri zenginliği sunsa da, genel eğilimleri görmek zor olabilir. Genel korelasyon değerlerinin özetlenmesi ve belki de en belirgin pozitif ve negatif ilişkilerin grafiksel olarak gösterilmesinin, bulguların etkisini artırabileceği düşüncesi ile örnek olarak olasılık oranı [likelihood ratio (LR+)] ve DOR arasındaki ilişkinin saçılım grafiği, [Şekil 2](#)'de verilmiştir.

[Şekil 2](#)'de düşük DOR değerlerinde LR+ değerlerinin daha sınırlı bir dağılım sergilediği, ancak DOR değerleri büyüdükçe LR+ değerlerinin de genel olarak artma eğiliminde olduğu görülmektedir.

TABLO 5: Örneklem büyüklüklerine göre DOR ile diğer kriterler arasındaki korelasyon değerleri

	GP	YN	YP	GN	Sens.	Spes.	PPV	NPV	LR+	LR-	Accur.	Youden
10	0,563"	-0,563"	-0,563"	0,563"	0,698"	0,698"	0,699"	0,698"	0,970"	-0,970"	0,949"	0,993"
15	0,558"	-0,558"	-0,558"	0,563"	0,696"	0,698"	0,696"	0,697"	0,968"	-0,967"	0,932"	0,990"
20	0,548"	-0,547"	-0,547"	0,533"	0,683"	0,687"	0,688"	0,685"	0,962"	-0,964"	0,916"	0,988"
25	0,539"	-0,543"	-0,535"	0,546"	0,677"	0,675"	0,686"	0,688"	0,963"	-0,962"	0,905"	0,986"
30	0,536"	-0,542"	-0,535"	0,563"	0,676"	0,687"	0,677"	0,690"	0,963"	-0,959"	0,903"	0,984"
35	0,506"	-0,540"	-0,521"	0,588"	0,676"	0,710"	0,650"	0,695"	0,963"	-0,961"	0,888"	0,986"
40	0,496"	-0,587"	-0,508"	0,552"	0,686"	0,685"	0,635"	0,713"	0,963"	-0,967"	0,885"	0,985"
45	0,558"	-0,485"	-0,588"	0,555"	0,668"	0,708"	0,722"	0,676"	0,961"	-0,961"	0,901"	0,984"
50	0,551"	-0,542"	-0,549"	0,573"	0,694"	0,709"	0,696"	0,715"	0,958"	-0,961"	0,901"	0,983"
55	0,528"	-0,542"	-0,572"	0,582"	0,688"	0,716"	0,698"	0,712"	0,963"	-0,963"	0,900"	0,984"
60	0,511"	-0,509"	-0,542"	0,572"	0,652"	0,704"	0,680"	0,689"	0,960"	-0,959"	0,888"	0,982"
65	0,524"	-0,492"	-0,557"	0,536"	0,657"	0,696"	0,698"	0,667"	0,959"	-0,954"	0,882"	0,979"
70	0,528"	-0,535"	-0,573"	0,573"	0,668"	0,716"	0,712"	0,711"	0,962"	-0,962"	0,904"	0,981"
75	0,583"	-0,558"	-0,556"	0,505"	0,708"	0,691"	0,710"	0,693"	0,962"	-0,961"	0,894"	0,982"
80	0,550"	-0,556"	-0,507"	0,564"	0,702"	0,703"	0,663"	0,700"	0,960"	-0,967"	0,898"	0,982"
85	0,591"	-0,521"	-0,504"	0,512"	0,709"	0,659"	0,696"	0,649"	0,957"	-0,958"	0,880"	0,978"
90	0,550"	-0,583"	-0,527"	0,558"	0,726"	0,695"	0,678"	0,701"	0,964"	-0,962"	0,891"	0,983"
95	0,577"	-0,554"	-0,529"	0,497"	0,719"	0,656"	0,703"	0,682"	0,959"	-0,951"	0,885"	0,976"
100	0,560"	-0,541"	-0,547"	0,542"	0,694"	0,713"	0,686"	0,674"	0,961"	-0,964"	0,887"	0,983"
125	0,518"	-0,575"	-0,507"	0,551"	0,702"	0,678"	0,656"	0,705"	0,961"	-0,961"	0,871"	0,980"
150	0,527"	-0,535"	-0,518"	0,541"	0,683"	0,701"	0,665"	0,680"	0,956"	-0,961"	0,881"	0,980"
175	0,559"	-0,519"	-0,543"	0,529"	0,681"	0,687"	0,703"	0,677"	0,960"	-0,960"	0,868"	0,980"
200	0,504"	-0,537"	-0,521"	0,549"	0,663"	0,684"	0,667"	0,697"	0,962"	-0,964"	0,871"	0,981"
225	0,518"	-0,506"	-0,573"	0,607"	0,651"	0,736"	0,702"	0,709"	0,959"	-0,959"	0,897"	0,975"
250	0,548"	-0,550"	-0,526"	0,527"	0,697"	0,680"	0,684"	0,687"	0,955"	-0,958"	0,877"	0,976"
300	0,518"	-0,582"	-0,519"	0,566"	0,702"	0,683"	0,674"	0,722"	0,957"	-0,955"	0,893"	0,976"
350	0,524"	-0,557"	-0,522"	0,540"	0,684"	0,686"	0,679"	0,691"	0,958"	-0,953"	0,870"	0,975"
400	0,562"	-0,519"	-0,504"	0,532"	0,695"	0,674"	0,680"	0,676"	0,958"	-0,956"	0,881"	0,976"
450	0,528"	-0,534"	-0,538"	0,549"	0,672"	0,690"	0,693"	0,688"	0,965"	-0,961"	0,870"	0,979"
500	0,547"	-0,467"	-0,512"	0,452"	0,663"	0,621"	0,689"	0,613"	0,958"	-0,953"	0,825"	0,972"
600	0,536"	-0,542"	-0,557"	0,548"	0,686"	0,701"	0,687"	0,699"	0,964"	-0,962"	0,890"	0,981"
700	0,554"	-0,524"	-0,534"	0,503"	0,692"	0,671"	0,690"	0,666"	0,957"	-0,959"	0,853"	0,976"
800	0,535"	-0,553"	-0,538"	0,495"	0,696"	0,665"	0,685"	0,684"	0,954"	-0,960"	0,869"	0,974"
900	0,543"	-0,519"	-0,516"	0,496"	0,672"	0,648"	0,676"	0,674"	0,958"	-0,962"	0,863"	0,977"
1.000	0,553"	-0,493"	-0,543"	0,503"	0,681"	0,686"	0,702"	0,636"	0,960"	-0,953"	0,864"	0,975"
Genel	0,364"	-0,357"	-0,357"	0,363"	0,686"	0,689"	0,686"	0,687"	0,960"	-0,960"	0,887"	0,981"

"Korelasyon $p < 0,001$ düzeyinde istatistiksel olarak anlamlı. GP: Gerçek pozitif; YN: Yalancı negatif; YP: Yalancı pozitif; GN: Gerçek negatif; Sens.: Duyarlılık; Spes.: Özgüllük; PÖD: Pozitif öngörü değeri; NÖD: Negatif öngörü değeri; LR: Olasılık oranı (likelihood ratio); Accur.: Doğruluk; Youden: Youden Index değeri



ŞEKİL 2: DOR değerleri ile LR+ değerlerinin değişimi

DOR: Tanısal olasılık oranı (diagnostic odds ratio)

TARTIŞMA

DOR, tanısal prosedürlerin ayırmacı gücünü genel olarak değerlendiren ve aynı zamanda 2 veya daha fazla tanı testinin doğruluğunu karşılaştırmak için kullanılan küresel bir tanısal doğruluk ölçüsüdür. Testin duyarlılığına ve özgüllüğüne büyük ölçüde bağlıdır. Yanlış pozitif ve yanlış negatif oranları düşük, duyarlılık ve özgüllük değerleri yüksek olan testler genellikle daha yüksek DOR değerine sahiptir. Testin duyarlılığı sabit kaldığında, özgüllük arttıkça DOR'un değeri de yükselir. Örneğin, duyarlılığı %90'ın üzerinde ve özgüllüğü %99 olan bir testin DOR değeri 500'den büyük olabilir. DOR, hastalığın yaygınlığından bağımsızdır. Ancak duyarlılık ve özgüllük gibi hastalığın ve çalışılan grubun patolojik durumlarının çeşitliliğine (hastalığın şiddeti, evresi, fazı, komorbiditeler vb.) bağlı olarak değişkenlik gösterebilir.¹⁰ DOR'un tersi, hasta olanlardaki negatiflik oranlarının, hasta olmayanlardaki negatiflik oranlarına oranı olarak yorumlanabilir. Duyarlılık veya özgüllük mükemmele yakın olduğunda DOR hızla yükselir. Dikkate alınması gereken bir diğer nokta, küresel bir ölçü olarak DOR'un belirli yaygınlıklarda bir testin hata oranlarını değerlendirmek için kullanılamayacağıdır. Aynı DOR'a sahip 2 test, farklı klinik sonuçlarla çok farklı duyarlılık ve özgüllüğe sahip olabilir.⁵ Çoğu tanı testinin birden fazla veya sürekli değişkenli sonuçları olmasına rağmen, bu sonuçları pozitif ya da negatif olarak sınıflandırmak için genellikle kategorilerin gruplanması veya belirli bir kesme noktası kullanılır. Bu göstergelerin hiçbiri tek başına bir testin ayırmacı performansını tam anlamıyla yansıtamaz. Duyarlılık, yalnızca ayırmacı kanıtın bir bileşenidir; çünkü yüksek duyarlılığa bazen düşük özgüllük eşlik edebilir. Ayrıca, duyarlılık ve özgüllüğü tek bir performans ölçüsü altında toplamak için basit bir aritmetik kural bulunmamaktadır. Bu nedenle, testin doğruluğu gibi kapsamlı bir performans göstergesine ihtiyaç duyulur. Test ile referans standardı arasındaki genel uyumun yanı sıra, doğruluk terimi özgül olarak testin hastaları doğru sınıflandırma oranını ifade eder. Bu oran, duyarlılık ve özgüllüğün eşit olmadığı durumlarda, çalışma grubundaki hastalığın yaygınlığına bağlıdır ve yanlış pozitif ile yanlış negatif sonuçları eşit şekilde değerlendirir. Buna ek olarak, tek bir performans göstergesi olarak kullanılan Youden İndeksi duyarlılık ve özgüllükten türetilir ve yaygınlıktan bağımsızdır. Ancak, duyarlılık ve özgüllüğün doğrusal bir dönüşümü olduğu için yorumlanması zor olabilir. Bu güçlüğüne rağmen Böhning ve ark. tarafından Youden İndeksi'nin DOR'a göre daha iyi bir tercih olabileceği ifade edilmektedir.¹² Test performansını tek bir değerle ifade etmek için kullanılan bir diğer metrik DOR'dur. Bu oran, yaygınlığa bağlı değildir ve epidemiyolojik bağlam-

da daha kolay anlaşılabilir.³ Tanısal olasılık oranı, test performansını değerlendiren bir ölçüt olarak, duyarlılık ve özgüllüğün prevalanstan bağımsız güçlü yönlerini, doğruluk avantajıyla tek bir değer altında toplar. Bu özellikleri sayesinde, yanlış negatif ve yanlış pozitif oranları arasındaki dengenin kritik olmadığı durumlarda testlerin karşılaştırılması için özellikle faydalıdır. Duyarlılık ve özgüllük, testin koşullu isabet oranlarını ifade ederken, tahmini değerler ve test sonrası olasılıklar, klinik uygulamalar açısından en anlamlı veriler arasında yer alır. Olasılık oranları, farklı test sonuçlarının tanısal değerlerini karşılaştırmada ve bu değerleri test sonrası olasılıklara dönüştürmede önemli bir araçtır. Bu göstergeler içinde, DOR hem köklü bir geçmişe hem de güçlü istatistiksel özelliklere sahip tek bir özet istatistik olarak öne çıkmaktadır.⁵ Günümüzde, uygulamalı makalelerde DOR'un diğer değerlerle birlikte raporlanması yaygın bir uygulama hâline gelmiştir. DOR'un ayrımcılık gücünü ölçen bir kriter olarak önerilmesi, uygulayıcıların onu bu şekilde yorumlamasını doğal kılmaktadır. Uygulamada DOR, optimum cut-off değerini belirlemek için bir ölçüt olarak kullanılmaktadır. DOR, 2'li sonuçlar ve test verileri için güvenilir bir değerlendirme sunar. Bununla birlikte verilerin sürekli olması durumunda DOR'un değeri, seçilen cut-off değerine bağlı olarak değişiklik gösterebilir.¹¹ Hasselblad ve Hedges tarafından, DOR'un cut-off değerinin geniş bir aralığında çok az değişime uğradığı belirtilmiştir.¹³ Bu, farklı çalışmalardan elde edilen sonuçların birleştirilmesi istendiğinde (metaanalizde olduğu gibi) yararlı bir özelliktir. Deeks, tanısal ve tarama testlerinin sistematik incelemeleri bağlamında, tanısal olasılık oranının genellikle tanısal eşikten bağımsız olarak makul ölçüde sabit, ancak ayrımcılık gücünü en üst düzeye çıkaran (hastalıklı ve hastalısız popülasyonları ayırmada) bir ölçü bulma bağlamında daha az avantajlı olduğunu belirtmektedir.¹⁴ DOR mutlak hastalık olasılığı hakkında hiçbir şey söylemez. Bu nedenle, epidemiyolojide önemli olmasına rağmen klinisyenler için daha az faydalıdır.¹⁵ Çalışmadan elde edilen sonuçlar, metriklerin tek başına kullanılmasının ve tanı ile ilgili araştırmalarda performans metriklerine ait tahminlerin doğruluğunu artırmak için, örneklem büyüklüğünü artırmanın yeterli olmayacağını, "doğru" denek seçmenin daha faydalı olacağını göstermiştir.

SONUÇ

DOR, tıbbi tanı testlerinin doğruluğunu ölçmede etkili ve özetleyici bir göstergedir. Testin performansını tek bir değerle ifade edebilmesi, özellikle tanı testlerinin karşılaştırılması ve metaanaliz çalışmalarında önemli bir referans noktası olmasını sağlamaktadır. DOR'un değerinin büyümesi, testin hem hastaları doğru şekilde tespit etme hem de sağlıklı bireyleri doğru şekilde dışlama konusunda başarılı olduğunu gösterir. Ancak, daha güvenilir sonuçlar elde edebilmek için tek başına değerlendirilmemeli, duyarlılık ve özgüllük gibi diğer metriklerle birlikte ele alınmalıdır.

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, bilirdişlik, herhangi bir firmada çalışma durumu, hissedarlık ve benzer durumları yoktur.

Yazar Katkıları

Bu çalışma hazırlanırken tüm yazarlar eşit katkı sağlamıştır.

KAYNAKLAR

1. Samawi H, Chen DG, Yin J, Alsharman M. Performance of diagnostic tests based on continuous bivariate markers. *J Appl Stat.* 2022;51(3):497-514. [\[Crossref\]](#) [\[PubMed\]](#) [\[PMC\]](#)
2. Pérez I, Taito-Vicenti IY, González-Xuriguera CG, Carvajal C, Franco JVA, Loézar C. How to interpret diagnostic tests. *Medwave.* 2021;21(7):e8432. Spanish, English. [\[Crossref\]](#) [\[PubMed\]](#)
3. Huang Y, Yin J, Samawi H. Methods improving the estimate of diagnostic odds ratio. *Commun Stat-Simul C.* 2018;47(2):353-66. [\[Link\]](#)
4. Casanova IJ, Campos M, Juarez JM, Gomariz A, Lorente-Ros M, Lorente JA. Using the diagnostic odds ratio to select patterns to build an interpretable pattern-based classifier in a clinical domain: multivariate sequential pattern mining study. *JMIR Med Inform.* 2022;10(8):e32319. [\[Crossref\]](#) [\[PubMed\]](#) [\[PMC\]](#)
5. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol.* 2003;56(11):1129-35. [\[Crossref\]](#) [\[PubMed\]](#)
6. Raslich MA, Markert RJ, Stutes SA. Selecting and interpreting diagnostic tests. *Biochem Med (Zagreb).* 2007;17(2):151-61. [\[Crossref\]](#)
7. Sousa MR, Ribeiro AL. Systematic review and meta-analysis of diagnostic and prognostic studies: a tutorial. *Arq Bras Cardiol.* 2009;92(3):229-38, 235-45. English, Spanish. [\[PubMed\]](#)
8. Zhou X, Obuchowski NA, McClish DK. *Statistical Methods in Diagnostic Medicine.* 2nd ed. Hoboken, New Jersey: John Wiley&Sons, Inc.; 2011. p. 238-9. [\[Crossref\]](#)
9. Lim CY. Methods for evaluating the accuracy of diagnostic tests. *Cardiovasc Prev Pharmacother.* 2021;3(1):15-20. [\[Crossref\]](#)
10. Šimundić AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC.* 2009;19(4):203-11. [\[PubMed\]](#) [\[PMC\]](#)
11. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med.* 2003;29(7):1043-51. [\[Crossref\]](#) [\[PubMed\]](#)
12. Böhning D, Holling H, Patilea V. A limitation of the diagnostic-odds ratio in determining an optimal cut-off value for a continuous diagnostic test. *Stat Methods Med Res.* 2011;20(5):541-50. [\[Crossref\]](#) [\[PubMed\]](#)
13. Hasselblad V, Hedges LV. Meta-analysis of screening and diagnostic tests. *Psychol Bull.* 1995;117(1):167-78. [\[Crossref\]](#) [\[PubMed\]](#)
14. Deeks JJ. Systematic reviews in health care: systematic reviews of evaluations of diagnostic and screening tests. *BMJ.* 2001;323(7305):157-62. [\[Crossref\]](#) [\[PubMed\]](#) [\[PMC\]](#)
15. Marill KA. Diagnostic and prognostic test assessment in emergency medicine: likelihood and diagnostic odds ratios. *Emerg Med J.* 2022;39(8):635-42. [\[Crossref\]](#) [\[PubMed\]](#)