

Tedavi İçin Gerekli Sayı İçin Güven Aralıklarının Güvenirliklerinin İncelenmesi: Bir Simülasyon Çalışması

Examining the Reliability of Confidence Intervals for the Number Needed to Treat: A Simulation Study

İsmet DOĞAN^a, Nurhan DOĞAN^a

^aAfyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim ABD, Afyonkarahisar, Türkiye

ÖZET Amaç: Literatürde, iki oran arasındaki fark için güven aralığının belirlenmesi amacıyla çok sayıda önerilmiş yöntem söz konusudur. Ancak spesifik olarak tedavi için gereken sayı [number needed to treat (NNT)] ile ilgili geliştirilmiş bir yöntem bulunmamaktadır. Dolayısıyla iki oran arasındaki fark için ileri sürülen yöntemlerden hangisinin NNT için kullanılabilmesinin belirlenmesi gerekmektedir. Bu çalışmanın amacı, Wald, Newcombe, Agresti-Caffo ve Anbar tarafından önerilen güven aralıklarının NNT için kullanılabilirliğini araştırmaktır. **Gereç ve Yöntemler:** Çalışmada, Python-random kütüphanesi kullanılarak $10 \leq n \leq 1000$ aralığında yer alan 35 farklı n değeri için veri türetilmiştir. Verilerin türetilmesinde önce a , b , c ve d ile gösterilen gözelerden hangisine değer atanacağı, sonra da ilgili göze atanacak değer belirlenmiştir. $n=10$ için 286, $n=15$ için 815 ve $n \geq 20$ için biner farklı veri seti çalışmada kullanılmıştır. **Bulgular:** Güven aralığının çarpıklığı ve güven aralığının NNT sayı değerini kapsaması kriterleri bakımından çalışmada dikkate alınan yöntemlerin iyiden kötüye doğru sıralaması Anbar, Wald, Agresti-Coffe ve Newcombe şeklindedir. Çalışmadan elde edilen bulgulara göre NNT ve NNT için önerilen güven aralığı örneklem büyüklüğünden etkilenmektedir. **Sonuç:** NNT ve NNT için önerilen güven aralığı hesaplama yöntemleri tatmin edici sonuçlar sunmamaktadır. Bu duruma, NNT sayı ve NNT için önerilen güven aralığı hesaplama yöntemlerinin sebep olduğu, dolayısıyla hesaplama ve yorumlamada mevcut zayıflıkları giderecek yeni yöntemlerin ortaya konulması gerektiği söylenebilir. Bütün bu olumsuzluklara rağmen güven aralığının çarpıklığı ve NNT sayı değerini kapsaması kriterleri bakımından Anbar yönteminin çalışmada dikkate alınan diğer yöntemlere tercih edilmesi gerektiği sonucuna ulaşılmıştır.

Anahtar kelimeler: Tedavi için gerekli sayı; güven aralığı; iki sonuçlu veri

ABSTRACT Objective: In the literature, there are many proposed methods for determining the confidence interval of the difference between 2 ratios. However, there is no developed method specifically for the number needed to treat (NNT). Therefore, it is necessary to determine which of the proposed methods for the differences between 2 ratios can be used for the NNT. The aim of this study is to investigate the usability of the confidence intervals proposed by Wald, Newcombe, Agresti-Caffo, and Anbar for the NNT. **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 Python-random library. In the derivation of the data, firstly, which cell shown with a , b , c and d will be assigned value, then the value to be assigned to the relevant cell was determined. 286 for $n=10$, 815 for $n=15$ and 1,000 different data sets for $n \geq 20$ were used in the study. **Results:** The order of the methods from best to worst is Anbar, Wald, Agresti-Coffe, and Newcombe. In ordering the methods, whether the confidence interval includes the NNT value and the skewness of the confidence interval were taken into account. According to the findings obtained from the study, the suggested confidence intervals for NNT and NNT are affected by the sample size. **Conclusion:** Calculation methods proposed for confidence interval NNT and NNT do not provide satisfactory results. It can be said that this situation is caused by the calculation methods suggested for the NNT confidence interval and NNT, so new methods should be introduced to eliminate existing weaknesses in calculation and interpretation. Despite all these negativities, it was concluded that the Anbar method should be preferred over the other methods considered in the study in terms of the skewness of the confidence interval and the criteria to include the NNT value.

Keywords: Number needed to treat; confidence 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 Türkiye Klinikleri Journal of Biostatistics.

Received: 11 Nov 2022 **Accepted:** 12 Dec 2022 **Available online:** 19 Dec 2022

2146-8877 / Copyright © 2023 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/>).

Tedavi için gereken sayı [the number needed to treat (NNT)] kavramından ilk olarak, Laupacis ve ark.nın tedavi sonuçlarına klinik olarak uygulanabilir ölçülerin ele alındığı dönüm noktası niteliğindeki makalelerinde bahsedilmiştir.¹ Çoğunlukla klinik denemelerde yeni geliştirilen bir tedavinin referans bir tedaviye ya da kontrole göre etkinliğini tespit etmek için kullanılmaktadır. NNT, yeni geliştirilen tedaviden ek bir hastada olumlu sonucun elde edilebilmesi için tedavi edilmesi gereken hasta sayısının bir tahminini sağlar. $NNT = 1$ yeni geliştirilen tedaviden olumlu bir sonuç gözlemek için bir hastayı tedavi etmenin yeterli olduğunu göstermekte iken, $NNT = k$ olması bir hastada yeni geliştirilen tedavinin olumlu etkisini görebilmek için ortalama olarak k tane hastanın yeni geliştirilen tedavi yöntemi ile tedavi edilmesi gerektiği anlamına gelmektedir.² $-\infty < NNT < -1$ olması, yeni geliştirilen tedavinin olumlu tedavi riskini artırdığını, $1 < NNT < \infty$ olması ise olumlu tedavi riskini azalttığını gösterir. $-1 < NNT < 1$ olması ise imkânsızdır. NNT, bir tedaviden elde edilen sonucun klinik ölçüsü olarak tedavi grubunda olumsuz bir olay (ters etki) riskini azaltmak için önerilmiş ve tıp alanına girmesinden bu yana popülerlik kazanmıştır.³ NNT, halk sağlığı ve tıbbi araştırmalarda yaygın olarak kullanılan bir etki ölçüsü olarak belirli bir dönemde tek bir kötü sonucu önlemek için gereken kişi sayısı olarak yorumlanır. Bu nedenle NNT, genellikle klinisyenlere ve politika yapıcılara bir tedavi veya tedavi politikasının etkisini tahmin etmede yardımcı olmayı amaçlar. Ancak gerçekte tedavi için ihtiyaç duyulan bir sayının uygun şekilde yorumlanması, gerçek etki idealize edilmiş etkiden farklı olacağı için düşünülen daha sorunlu olabilir.⁴ Matematiksel olarak risk farklarıyla ilişkili olan NNT, tedavi edilen hastalar açısından yorumlamayı kolaylaştırdığından, terapötik karar verme ve hasta başı öğretim için bir araç olarak yaygın olarak kullanılmaktadır.⁵ NNT, istatistiksel bir kavramdır, ancak sezgiseldir. Orijinal hâliyle NNT, olumsuz bir olay için 2 tedavinin karşılaştırıldığı durumlarda insidans oranlarına dayalı sağlık hizmeti müdahalesinin etkinliğini gösteren basit epidemiyolojik bir ölçüdür. Randomize kontrollü çalışmalar, tedavinin faydalı ve zararlı etkilerini belirlemek için altın standart olarak kabul edilir. Tıbbın ilk ilkelerinden biri “önce zarar verme” olduğundan, tedavinin yararları düşünüldüğünde, klinisyenler riskleri de incelemelidir.⁶ Hastalar, klinisyenler ve sağlık alanı ile ilgili politika yapıcılar tarafından verilen kararlara rehberlik etmek için kullanılan kontrollü klinik araştırmalarda, genellikle yalnızca göreceli etki ölçülerinin değerleri verilmektedir. Oysa mutlak risk azaltma [absolute risk reduction (ARR)], NNT ve zarar vermek için gereken sayı [number needed to harm (NNH)] gibi mutlak ölçüler, daha kolay yorumlanabilir, klinik olarak daha anlamlı olabilir ve sonuç riski düşük olduğunda farklılıkları abartma olasılığı daha düşüktür.⁷ Rölatif risk, ARR, odds oranı ve hazard oranı gibi tedavi ve kontrol gruplarının karşılaştırılmasında kullanılan istatistiksel ölçüler olmasına rağmen tedavi prosedürlerinin hem istatistiksel hem de klinik öneminin ortaya konulmasında doğrudan bir karşılaştırma sağlayan NNT, 2 sonuçlu (başarı/başarısızlık, evet/hayır, komplikasyon var/komplikasyon yok vb.) randomize kontrollü çalışmaların sonuçlarından elde edilen ve yaygın olarak kullanılan bir ölçüdür.⁸ Müdahalenin sonuçlarını eylemsizlikle karşılaştırma, müdahalelerle ilişkili potansiyel zararları veya olumsuz olayları ölçme ve bir müdahalenin teşhis, tedavi veya durumların önlenmesindeki yansımalarının diğer müdahalelere karşı değerlendirilebileceği durumlarda, tedavi yöntemlerinin karşılaştırılmasını sağladığı için NNT'nin diğer istatistiksel ölçülere göre daha kullanılabilir olduğu belirlenmiştir. Bir klinik araştırma tamamlandığında, tedavi ve kontrol gruplarında istenen sonucun gözlemlenme oranları rapor edilir. NNT bu değerlerden türetilir ve tedavinin gözlemlenen hastalık üzerindeki etkisinin büyüklüğünü gösterir. Tedavi ve kontrol grupları arasındaki mutlak risk azalmasının karşılığı alınarak hesaplanır. Riskler, bilimsel literatürde, tedavi etkisinin veya zararının ölçütü olarak rutin olarak rapor edilmektedir. Bir tedavinin diğerine göre yararını vurgulamada veya tedavinin uygulanabilir bir seçenek olup olmadığını belirlemede çok önemlidirler.⁹ Önceki tedavi ile yeni tedavinin karşılaştırıldığı bir çalışmada, araştırmacılar tarafından cevaplanması istenen klinik sorulardan biri, yeni tedavinin faydalı etkisinin olup olmadığı veya hastalara zarar verip vermediğidir. Üstelik araştırmacılar, önceki tedavi ile yeni tedavi arasındaki iyileştirme farkının sayısal büyüklüğüne ve farkın anlamlı olup olmadığına da odaklanmaktadırlar. Bu soruları cevaplamak için etki büyüklüğünü ölçen standartlaştırılmış ortalama fark, rölatif risk, odds oranı, NNT vb. kriterler kullanılmaktadır.¹⁰ NNT, çalışma tamamlandıktan sonra ilave bir iyi sonuç elde etmek için ortalama olarak kaç hastanın tedavi edilmesi gerektiğini gösterir ve hem terapötik hem de tanısal çalışmalardan elde edilen birçok sonuç türüne uygulanabilir. Bir tedavi arzu edilen sonuçları artırdığında, ortaya çıkan

değer, fayda sağlamak için gereken sayıdır. Bir tedavi advers olayları artırdığında ortaya çıkan değer, NNH'dir. NNH teriminden kaçınmak için Altman, NNT yerine ek yararlı bir sonuç için tedavi edilmesi gereken sayı [number needed to treat for an additional beneficial outcome (NNTB)] ve NNH yerine ek zararlı bir sonuç için tedavi edilmesi gereken sayı [number needed to treat for an additional harmful outcome (NNTH)] terimlerini önermiştir. Etki yönü bakımından NNTB'nin değerinin yüksek olması, tedavinin az etkili olduğunu, NNTH değerinin yüksek olması ise tedavinin güvenilir olduğunu göstermektedir.^{11,12} Tanı stratejilerine uygulandığında, ortaya çıkan değerler asemptomatik bireylerde testleri taramak için semptomatik bireylerde ise tanı koymak için gereken sayıyı ifade etmektedir.¹³ Genellikle 2 sonuçlu randomize klinik çalışmalarda, etkinliği ölçmede kullanılan NNT için mevcut 3 farklı yaklaşım söz konusudur. Furukawa-Leucht's ve Kraemer-Kupfer's tarafından ortaya konulan yaklaşımlar, Cohen'in *d* indeksini NNT'ye dönüştürmeye odaklanırken, Laupacis ve ark.nın yaklaşımı reformülasyondan ziyade öncelikle bu sayının tahmini ile ilgilenir. Laupacis ve ark.nın yaklaşımının, Kraemer-Kupfer'in yaklaşımından özünde farklı olduğu, Furukawa-Leucht yaklaşımının yalnızca eşit standart sapmalara sahip normal dağılımlı sonuçlar için kullanılmaya uygun olduğu gösterilmiştir.¹⁴⁻¹⁶ Ancak çalışma sonuçlarını özetleyen tek bir ölçünün varlığı kolaylık gibi görünmesine rağmen tek bir sayıya indirgenmiş çalışma sonuçlarının eksikliklerinin olması kaçınılmazdır. Yararlı bir ölçü olduğu düşünülse de NNT sorunsuz değildir. Yanlış yorumlama ve yanlış kullanım potansiyeline sahiptir. Alternatif tedavi, tedavi süresi, takip süresi ve etkinin yönü belirtilmeden bir NNT yorumlanamaz. NNT'nin hesaplanması için kişi-zaman verilerinin kullanılması genellikle uygun değildir. Özellikle hesaplama sonucunda değeri tamsayı olarak elde edilemeyen NNT değerlerinin kabaca bir sonraki daha yüksek tam sayıya yuvarlanması tedaviler arasındaki farklılıkları gizleyebilir. NNT'nin kullanımı ve yorumlanması;

- Müdahaleyi uygulamanın ne kadar zor olduğuna ve müdahalenin ne kadar erişilebilir ve uygun maliyetli olduğuna bağlı olabileceğinden, klinik içgörü, hasta değerleri ve koşulları,

- Takip süresi (farklı takip sürelerine sahip çalışmalar için NNT'ler doğrudan karşılaştırılabilir olmadığından),
- Olayın temel riski,
- NNT'nin istatistiksel özellikleri,
- Müdahalenin karşılaştırıldığı alternatif tedavi,
- Sonuç,
- Etki ölçüsünün yönü ve büyüklüğü,
- NNTB ve NNTH ölçeği,
- NNTB veya NNTH'ye ait güven aralıkları

faktörlerinin anlaşılmasını gerektirir.^{3,12,17} Vaka sayısının azaltılması gerektiğini gösteren negatif değerler veya vaka sayısının artması gerektiğini gösteren pozitif değerler daha az kafa karışıklığına sebep olmasına rağmen NNT'nin kendisi bir *p* – değeri veya güven aralığı ile birlikte sunulmadıkça istatistiksel anlamlılık taşımaz. NNT'ler için güven aralıkları genellikle istatistiksel olarak anlamlı sonuçlar için verilir. Bunun başlıca nedeni, $ARR = 0$ olduğunda belirsizliğin olmasıdır.^{12,13} NNT'ler için belirsizlik bir güven aralığı ile değerlendirilmelidir. En sık kullanılan güven aralığı %95 güven aralığıdır. NNT için %95 güven aralığı, NNT'nin kendisine benzer şekilde türetilir. ARR'nin %95 güven aralığı için elde edilen sınırların karşılığı türetilir ve daha sonra tersine çevrilerek NNT için %95 güven aralığının sınırları verilir. NNT için tahmin edilen değer, prognostik olarak çalışma katılımcılarına benzer bir popülasyonda tedavi edilmesi gereken sayının en olası değerini temsil eden bir "nokta tahmini" olarak kabul edilir. Bununla birlikte, herhangi bir istatistikte olduğu gibi nokta tahmininin kesinliğini, popülasyondaki istatistiğin gerçek değeri olarak belgelemek önemlidir. Kesinliği ifade etmek için en yaygın yöntem, istatistiğin gerçek değerinin içinde yer alması beklenen bir üst aralığı ve alt aralığı temsil eden güven aralıklarını bildirmektir. Tedavi edilmesi gereken sayı için %95 güven aralığı, tedavi edilmesi gereken gerçek sayının zamanın %95'inde beklendiği üst ve alt aralığı temsil eder. Belirgin bir mutlak risk azalmasıyla gruplar arasındaki farklılıklardaki varyansın birincil kaynağı olarak şansın işlemediği varsayılabi-

li; bu nedenle NNT ve ilişkili güven aralıkları klinik kararlar vermek için yararlıdır. ARR için güven aralıklarını hesaplamada en yaygın olarak Wald yöntemi kullanılmaktadır. Örnek boyutunun büyük olması ve ARR'nin 0 veya 1'e yakın olmaması durumunda güven aralıklarını belirlemek için yeterli bir yöntemdir. Daha çok tercih edilen ise Wilson yöntemidir.¹⁸ Wald güven aralıkları, geniş örneklem büyüklükleri için yeterli olmakla birlikte, pek çok pratik durumda zayıf kapsama özelliklerine ve sapma eğilimine sahiptirler. Özellikle küçük örneklerde ve dengesiz tasarımlarda Wald yöntemi güvenilmez, hatta teorik olarak imkânsız sonuçlara yol açmaktadır.¹⁹ NNT bir etki ölçüsü olduğu için risk oranı, odds oranı ve risk farkı gibi diğer etki ölçülerine benzer şekilde güven aralıkları hesaplanmalıdır.³ Bu çalışmanın amacı, Wald, Newcombe, Agresti-Caffo ve Anbar tarafından önerilen güven aralıklarının NNT için kullanılabilirliğini araştırmaktır. Makalede Helsinki Deklarasyonu prensipleri dikkate alınmıştır.

GEREÇ VE YÖNTEMLER

Çalışmada, 2×2 tablolardan elde edilen NNT ve güven aralıkları 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, Python-random kütüphanesi kullanılarak $10 \leq n \leq 1000$ 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, 1000) için veri türetilmiştir. Veri türetimine ilişkin Python 3.9.10 (Python Software Foundation, ABD) programlama dili 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' ncü 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 ARR ve NNT değerleri hesaplanır.

NNT, hesaplama yönteminin seçimi, çalışmanın tasarımı ve ilgilenilen olayı tanımlayan değişkenlerin tipine bağlı olacaktır. NNT, ilgilenilen olay için elde edilen veriler 2 sonuca sahipse hesaplanabilir, sürekli veriler için hesaplanamaz.²⁰ ARR ve NNT'nin hesaplanmasında kullanılan formüller dikkate alındığında, klasik 2×2 tablo örneği [Tablo 1](#)'de verildiği şekilde oluşturulmaktadır.

TABLO 1: Hesaplama tablosu.

	Olumsuz sonuç	Olumlu sonuç	Toplam
Deney grubu	<i>a</i>	<i>b</i>	<i>n</i> ₁
Kontrol grubu	<i>c</i>	<i>d</i>	<i>n</i> ₂

ARR ve NNT'nin değeri *a, b, c ve d* frekanslarına bağlıdır. *a, b, c ve d*'ye dayalı ARR'nin değeri Eşitlik 1-2, NNT'nin değeri ise Eşitlik 3-4 kullanılarak elde edilmektedir.

İlgilenilen olay olumsuz ise ARR;

$$ARR = \frac{c}{c+d} - \frac{a}{a+b} = \hat{p}_1 - \hat{p}_2 \quad (1)$$

İlgilenilen olay olumlu ise ARR;

$$ARR = \frac{a}{a+b} - \frac{c}{c+d} = \hat{p}_2 - \hat{p}_1 \quad (2)$$

eşitlikleri kullanılarak hesaplanmaktadır. ARR, yeterli bir istatistiktir, ancak klinik araştırmanın risk sonucunu hastaların kolayca anlayabileceği şekilde gösteren bir ölçü değildir. Üstelik bir klinik deneyde rapor edilen ARR etkileyici görünse de klinisyenlere NNT hakkında doğrudan kullanılabilir bilgi sağlamaz.¹⁸ Bazı tedaviler için ARR, belirli bir advers olay açısından orta ya da yüksek risk taşıyan hastalar için istatistiksel olarak anlamlı ve klinik olarak önemli olabilir. Ancak düşük risk taşıyan hastalar için aktif tedavinin toksisitesini ve maliyetini garanti etmede yeterli olmayabilir.⁵ ARR'nin değerinin oran olarak bildirilmesi durumunda NNT'nin değeri,

$$NNT = \frac{1}{ARR} \quad (3)$$

eşitliği kullanılarak, yüzde olarak bildirilmesi durumunda ise

$$NNT = \frac{100}{ARR} \quad (4)$$

eşitliği kullanılarak hesaplanır. Eşitlik 3-4'ten elde edilen değerlerin ondalıklı olarak elde edilmesi durumunda NNT en yakın tam sayıya yuvarlanır. Tedavi ne kadar etkili olursa ARR o kadar büyük, NNT ise o kadar küçük olur. $NNT = 1$ olması, yeni tedavinin uygulandığı tüm hastalarda olumsuz bir sonucun ortaya çıkmasının mümkün olmadığını gösterir. Yeni tedavinin ilgilenilen olay üzerinde yararlı bir etkisinin olmaması durumunda ise $ARR = 0$ ve $NNT = \infty$ olur. Bu durum, yeni tedavi yöntemi ile sonsuz sayıda hasta tedavi edilse bile, tedavi üzerinde hiçbir yararlı etkisinin olmayacağını, kontrol grubunun genel olarak daha iyi sonuçlara sahip olduğunu gösterir. $NNT < 0$ değeri klinisyene, yeni tedavinin hastalara zarar verdiğini düşünmeye sevk etmelidir. Bu durumda, NNT'nin aksine NNH'den bahsedilir. $ARR \leq 0$ olması NNT için güven aralığı hesaplamalarında sorun yarattığı için istenmeyen bir durumdur.^{18,20} NNT için güven aralığı, ARR için elde edilen güven sınırlarının ters çevrilmesi ve değiştirilmesiyle elde edilebilir. ARR için alt ve üst güven sınırları sırasıyla LL_{ARR} ve UL_{ARR} olarak ifade edildiğinde, NNT için alt sınır $LL_{NNT} = 1/UL_{ARR}$ ve üst sınır $UL_{NNT} = 1/LL_{ARR}$ eşitlikleri kullanılarak hesaplanır.¹⁹ ARR hem pozitif hem de negatif değerler alabildiğinden, logaritmik dönüşüme ihtiyaç duymamaktadır. ARR için %95 güven aralığı,

$$ARR \mp 1.96 * \sqrt{Var(ARR)} \quad (5)$$

eşitliği kullanılarak hesaplanmaktadır. Güven aralığının "0" değerini içermemesi ARR'nin istatistiksel olarak anlamlı olduğunu göstermektedir.²¹ Güven aralıklarının hesaplanmasında kullanılan çok sayıda yöntem bulunmaktadır. Bu çalışmada, dikkate alınan güven aralıklarının hesaplanması ile ilgili formüller Eşitlik 6-9 ile verilmiştir. Alt ve üst sınırların hesaplanması Excel uygulaması ile gerçekleştirilmiştir. Eşitliklerde yer alan

$c(\alpha)$ ve $z_{\alpha/2}$ 'nin değeri %95 güven aralığı için 1,96 olarak alınmıştır. ARR için güven aralıklarının hesaplanmasında Eşitlik 1 ile verilen durum göz önünde bulundurulmuştur.

Wald Yöntemi için güven aralığı²²

$$ARR \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}} \quad (6)$$

Newcombe Hibrid Skor Yöntemi için güven aralığı²¹

$$\hat{p}_i \left(\frac{n_i}{n_i + z_{\alpha/2}^2} \right) + \frac{1}{2} \left(\frac{z_{\alpha/2}^2}{n_i + z_{\alpha/2}^2} \right) \pm z_{\alpha/2}^2 \sqrt{\frac{1}{n_i + z_{\alpha/2}^2} \left[\hat{p}_i(1 - \hat{p}_i) \left(\frac{n_i}{n_i + z_{\alpha/2}^2} \right) + \frac{1}{4} \left(\frac{z_{\alpha/2}^2}{n_i + z_{\alpha/2}^2} \right) \right]} \quad i = 1, 2$$

l_1 : $i = 1$ için alt sınırlar,

l_2 : $i = 2$ için alt sınırlar,

u_1 : $i = 1$ için üst sınırlar,

u_2 : $i = 2$ için üst sınırlar,

$$\begin{cases} \text{Alt Sınır} = ARR - \sqrt{(\hat{p}_1 - l_1)^2 + (u_2 - \hat{p}_2)^2} \\ \text{Üst Sınır} = ARR + \sqrt{(\hat{p}_2 - l_2)^2 + (u_1 - \hat{p}_1)^2} \end{cases} \quad (7)$$

Anbar Yöntemi için güven aralığı²³

$$\bar{\delta} = \frac{ARR}{1 + c(\alpha)/n_2} - \frac{\left(\hat{p}_1 - \frac{1}{2}\right) c(\alpha)/n_2}{1 + c(\alpha)/n_2}$$

$$L = \sqrt{\frac{c(\alpha)/n_2}{[1 + c(\alpha)/n_2]^2} \left[(1 + c(\alpha)/n_2) \left\{ \frac{1}{4} \left(1 + \frac{n_2}{n_1} \right) - \frac{n_2}{n_1} \left(\hat{p}_1 - \frac{1}{2} \right)^2 \right\} - \left(\frac{1}{2} - \hat{p}_2 \right)^2 \right]}$$

$$\begin{cases} \text{Alt Sınır} = \bar{\delta} - L \\ \text{Üst Sınır} = \bar{\delta} + L \end{cases} \quad (8)$$

Agresti-Caffo Yöntemi için güven aralığı²⁴

$$\check{p}_1 - \check{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\check{p}_1(1 - \check{p}_1)}{\check{n}_1} + \frac{\check{p}_2(1 - \check{p}_2)}{\check{n}_2}} \quad (9)$$

$$\check{n}_1 = n_1 + 2$$

$$\check{n}_2 = n_2 + 2$$

$$\check{p}_1 = \frac{a + 1}{\check{n}_1}$$

$$\check{p}_2 = \frac{c + 1}{\check{n}_2}$$

NNT'nin hesaplandığı bir çalışmada, eğer tedavinin bir etkisi yoksa ARR'nin değeri 0, NNT'nin değeri ise sonsuz olacaktır. Tedavi grubuna ait oran kontrol grubuna ait orandan daha düşük olduğu durumda, güven aralığının alt sınırı negatif, üst sınırı ise pozitif olacaktır. NNT için fayda, zarar ve sonsuzluğu içerebilen güven aralıkları oldukça kafa karıştırıcıdır.²⁵ Bütün bu matematiksel tuhaflıklara rağmen güven aralıkları, bulguların genel güvenilirliğini gösterdiğinden, tüm tahminlerde olduğu gibi NNT tahmini için de önemlidir. Güven aralığının yalnızca sonucun anlamlı olduğu durumlar için elde edilmesi düşüncesi, bir denemenin sonucu önemli olmadığında güven aralığının özellikle yararlı olduğu yönündeki tavsiyeye aykırı olduğundan tatmin edici değildir. Tedavi etkisi %5 düzeyinde anlamlı olduğunda, ARR için %95 güven aralığı 0 değerini dolayısıyla NNT için %95 güven aralığı da sonsuzu içermeyecektir.¹¹ NNT ile birlikte %95'lik güven aralığının verilmesi, yeni önerilen tedaviden sağlanacak faydadaki belirsizliğin de etkili bir şekilde ortaya konulmasını sağlar. Tedavi etkisine ilişkin diğer göstergeler bu amaç için pek uygun değildir. Örneğin istatistiksel olarak anlamlı bir p değeri, klinik önemden ziyade istatistiksel önem taşır. p değeri, tedavi seçimiyle ilişkili sonuçlarda bir fark olacağını gösterir, ancak bu farkın ne kadar büyük olacağını göstermez.¹³

BULGULAR

Randomize kontrollü çalışmalar, geleneksel tedaviye kıyasla yeni bir tedavinin etkisinin en geçerli tahminini sağlar. Klinik kanıtları bireysel hastaların bakımına uygulamak isteyen klinisyenlerin, klinik araştırmalarda ve sistematik incelemelerde sunulan verilerden ARR ve NNT'yi anlaması ve hesaplayabilmesi gerekir.²⁶ ARR, tedavi ve sonuç arasındaki ilişkinin gerçek büyüklüğünün belirlenmesinde yaygın olarak kullanılan geleneksel ölçülerden biridir. Hem temel riski hem de risk azalmasının büyüklüğünü hesaba kattığı için klinik karar verme açısından tedavi etkisi ile en çok alakası olan ölçüdür.²⁷ ARR için çalışmada elde edilen sonuçlar [Tablo 2](#)'de verilmiştir.

[Tablo 2](#)'den de görüldüğü üzere, örneklem büyüklüğü arttıkça belirsiz ya da 0 değerine sahip ARR sayısı azalmaktadır. Özellikle $n > 60$ olması durumunda ARR ile ilgili olumsuzluklar ortadan kalkmaktadır. ARR değerinin belirsiz ya da 0 olmadığı durumlarda ise örneklem büyüklüğüne bağlı olmaksızın ARR için elde edilecek değerlerin yarısının negatif, yarısının da pozitif olacağı neredeyse kesin bir şekilde ifade edilebilir. ARR'nin istatistiksel olarak anlamlı olup olmaması üzerinde örneklem büyüklüğünün ve çalışmada dikkate alınan güven aralığı hesaplama yöntemlerinin etkileri incelenmiş, elde edilen sonuçlar [Tablo 3](#)'te verilmiştir.

[Tablo 3](#)'ten de görüldüğü üzere, örneklem büyüklüğü arttıkça "0" değerini içermeyen güven aralığı sayısı azalmaktadır. Çalışmada dikkate alınan yöntemlerin tamamında $n \leq 60$ olması durumunda 0 değerini içermeyen güven aralığı sayısı %60 ve daha yüksek değerlere sahip olmaktadır. Sıfır değerini içermeyen güven aralığı sayısı bakımından çalışmada dikkate alınan yöntemler iyiden kötüye doğru Newcombe, Agresti-Caffo, Wald ve Anbar şeklinde sıralanmaktadır. Özellikle $n \leq 40$ olduğu durumlarda Newcombe tarafından önerilen yöntem 0 değerini içermeyen güven aralığı sayısının %90 ve daha yüksek değerlere sahip olmasından dolayı diğer yöntemlere göre öne çıkmaktadır. NNT, tedavilerin etkinliğini sıralamak için istatistiksel olarak sağlam ve kolayca yorumlanabilir bir ölçü olarak geliştirilmiş olmasına rağmen hesaplama ve yorumlamada belirli zayıflıklara sahiptir. $NNT = 1$, $NNT = \infty$ ya da $NNT < 0$ olması yeni tedavinin faydalı olmadığını göstermektedir. Üstelik $NNT = \infty$ ve $ARR \leq 0$ olması NNT için güven aralığı hesaplamalarında sorun yarattığı için istenmeyen durumlardır. Bu durumlar ile ilgili çalışmadan elde edilen sonuçlar örneklem büyüklüklerine göre incelenmiş ve [Tablo 4](#)'de verilmiştir.

[Tablo 4](#)'ten de görüldüğü üzere $n < 30$ olması durumunda NNT'nin aksine NNH'den bahsedilir. $n \geq 30$ olması durumunda ise sonuçların yarısının NNT'yi, yarısının da NNH'yi göstereceği neredeyse kesin bir şekilde ifade edilebilir. NNT'nin istatistiksel özellikleri sorunludur ve kullanımının bazı dezavantajları bulunmaktadır. Özellikle küçük örneklem büyüklüklerinin söz konusu olduğu durumlarda, ARR için elde edilen güven aralıkları NNT için güven aralığına dönüştürüldüğünde çok büyük bir etkiye sahip olabilecek kötü sonuçlar verebilir. Önemsiz bir NNT değeri, biri tedavinin gerçekten zararlı olma olasılığına diğeri ise

tedavinin faydalı olma olasılığına izin veren 2 bölümden oluşan bir güven aralığına sahip olabilir.²⁸ Yöntemlere göre alt sınır değerinin üst sınır değerinden büyük olması ile ilgili çalışmadan elde edilen sonuçlar örneklem büyüklüklerine göre incelenmiş ve [Tablo 5](#)'te verilmiştir.

TABLO 2: ARR ile ilgili istatistikler.

n	Wald, Newcombe ve Anbar yöntemleri için				Agresti-Caffo yöntemi için		
	Belirsiz	Sıfır	Negatif	Pozitif	Sıfır	Negatif	Pozitif
10	22 (%7,7)	26 (%9,1)	119 (%41,6)	119 (%41,6)	12 (%4,2)	137 (%47,9)	137 (%47,9)
15	31 (%3,8)	44 (%5,4)	370 (%45,4)	370 (%45,4)	-	407 (%49,9)	408 (%50,1)
20	23 (%2,3)	39 (%3,9)	471 (%47,1)	467 (%46,7)	25 (%2,5)	492 (%49,2)	483 (%48,3)
25	15 (%1,5)	27 (%2,7)	488 (%48,8)	470 (%47,0)	-	512 (%51,2)	488 (%48,8)
30	9 (%0,9)	26 (%2,6)	461 (%46,1)	504 (%50,4)	4 (%0,4)	477 (%47,7)	519 (%51,9)
35	7 (%0,7)	12 (%1,2)	486 (%48,6)	495 (%49,5)	8 (%0,8)	486 (%48,6)	506 (%50,6)
40	7 (%0,7)	14 (%1,4)	490 (%49,0)	489 (%48,9)	5 (%0,5)	498 (%49,8)	497 (%49,7)
45	4 (%0,4)	11 (%1,1)	503 (%50,3)	482 (%48,2)	5 (%0,5)	512 (%51,2)	483 (%48,3)
50	2 (%0,2)	9 (%0,9)	489 (%48,9)	500 (%50,0)	6 (%0,6)	495 (%49,5)	499 (%49,9)
55	1 (%0,1)	8 (%0,8)	497 (%49,7)	494 (%49,4)	-	506 (%50,6)	494 (%49,4)
60	-	13 (%1,3)	493 (%49,3)	494 (%49,4)	-	502 (%50,2)	498 (%49,8)
65	1 (%0,1)	5 (%0,5)	496 (%49,6)	498 (%49,8)	2 (%0,2)	491 (%49,1)	507 (%50,7)
70	2 (%0,2)	8 (%0,8)	480 (%48,0)	510 (%51,0)	4 (%0,4)	482 (%48,2)	514 (%51,4)
75	2 (%0,2)	4 (%0,4)	507 (%50,7)	487 (%48,7)	-	509 (%50,9)	491 (%49,1)
80	1 (%0,1)	3 (%0,3)	485 (%48,5)	511 (%51,1)	2 (%0,2)	489 (%48,9)	509 (%50,9)
85	1 (%0,1)	-	513 (%51,3)	486 (%48,6)	-	511 (%51,1)	489 (%48,9)
90	1 (%0,1)	2 (%0,2)	511 (%51,1)	486 (%48,6)	-	519 (%51,9)	481 (%48,1)
95	2 (%0,2)	4 (%0,4)	509 (%50,9)	485 (%48,5)	-	512 (%51,2)	488 (%48,8)
100	2 (%0,2)	3 (%0,3)	465 (%46,5)	530 (%53,0)	2 (%0,2)	471 (%47,1)	527 (%52,7)
125	-	1 (%0,1)	492 (%49,2)	507 (%50,7)	-	495 (%49,5)	505 (%50,5)
150	-	-	500 (%50,0)	500 (%50,0)	-	497 (%49,7)	503 (%50,3)
175	-	1 (%0,1)	477 (%47,7)	522 (%52,2)	-	478 (%47,8)	522 (%52,2)
200	-	2 (%0,2)	513 (%51,3)	485 (%48,5)	1 (%0,1)	511 (%51,1)	488 (%48,8)
225	-	2 (%0,2)	485 (%48,5)	513 (%51,3)	-	484 (%48,4)	516 (%51,6)
250	-	-	480 (%48,0)	520 (%52,0)	-	483 (%48,3)	517 (%51,7)
300	-	-	487 (%48,7)	513 (%51,3)	-	488 (%48,8)	512 (%51,2)
350	-	-	486 (%48,6)	514 (%51,4)	1 (%0,1)	484 (%48,4)	515 (%51,5)
400	-	-	503 (%50,3)	497 (%49,7)	-	504 (%50,4)	496 (%49,6)
450	-	-	523 (%52,3)	477 (%47,7)	-	525 (%52,5)	475 (%47,5)
500	-	-	501 (%50,1)	499 (%49,9)	1 (%0,1)	499 (%49,9)	500 (%50,0)
600	-	-	527 (%52,7)	473 (%47,3)	-	525 (%52,5)	475 (%47,5)
700	-	-	502 (%50,2)	498 (%49,8)	-	500 (%50,0)	500 (%50,0)
800	-	-	522 (%52,2)	478 (%47,8)	-	522 (%52,2)	478 (%47,8)
900	-	-	495 (%49,5)	505 (%50,5)	-	495 (%49,5)	505 (%50,5)
1.000	-	-	500 (%50,0)	500 (%50,0)	-	500 (%50,0)	500 (%50,0)
Genel	133 (%0,4)	264 (%0,8)	16826 (%49,3)	16878 (%49,5)	78 (%0,2)	16998 (%49,8)	17025 (%49,9)

ARR: Mutlak risk azaltma.

TABLO 3: ARR için yöntemlere göre sıfır değerini içermeyen güven aralığı sayısı.

n	Wald	Newcombe	Agresti-Caffo	Anbar
10	196 (%68,5)	163 (%57,0)	257 (%89,9)	172 (%60,1)
15	592 (%72,6)	554 (%68,0)	707 (%86,7)	509 (%62,5)
20	744 (%74,4)	757 (%75,7)	851 (%85,1)	624 (%62,4)
25	746 (%74,6)	799 (%79,9)	827 (%82,7)	639 (%63,9)
30	710 (%71,0)	850 (%85,0)	776 (%77,6)	605 (%60,5)
35	710 (%71,0)	855 (%85,5)	775 (%77,5)	613 (%61,3)
40	721 (%72,1)	861 (%86,1)	779 (%77,9)	625 (%62,5)
45	715 (%71,5)	860 (%86,0)	746 (%74,6)	623 (%62,3)
50	700 (%70,0)	851 (%85,1)	727 (%72,7)	598 (%59,8)
55	710 (%71,0)	842 (%84,2)	732 (%73,2)	615 (%61,5)
60	708 (%70,8)	850 (%85,0)	731 (%73,1)	613 (%61,3)
65	684 (%68,4)	860 (%86,0)	706 (%70,6)	603 (%60,3)
70	677 (%67,7)	838 (%83,8)	691 (%69,1)	586 (%58,6)
75	679 (%67,9)	816 (%81,6)	694 (%69,4)	594 (%59,4)
80	676 (%67,6)	824 (%82,4)	688 (%68,8)	594 (%59,4)
85	675 (%67,5)	820 (%82,0)	690 (%69,0)	595 (%59,5)
90	676 (%67,6)	800 (%80,0)	687 (%68,7)	594 (%59,4)
95	652 (%65,2)	798 (%79,8)	671 (%67,1)	586 (%58,6)
100	640 (%64,0)	803 (%80,3)	656 (%65,6)	556 (%55,6)
125	654 (%65,4)	782 (%78,2)	664 (%66,4)	578 (%57,8)
150	629 (%62,9)	752 (%75,2)	640 (%64,0)	558 (%55,8)
175	627 (%62,7)	754 (%75,4)	629 (%62,9)	549 (%54,9)
200	650 (%65,0)	723 (%72,3)	658 (%65,8)	580 (%58,0)
225	599 (%59,9)	718 (%71,8)	603 (%60,3)	536 (%53,6)
250	592 (%59,2)	712 (%71,2)	595 (%59,5)	524 (%52,4)
300	581 (%58,1)	690 (%69,0)	583 (%58,3)	529 (%52,9)
350	573 (%57,3)	670 (%67,0)	575 (%57,5)	538 (%53,8)
400	599 (%59,9)	658 (%65,8)	603 (%60,3)	549 (%54,9)
450	605 (%60,5)	655 (%65,5)	606 (%60,6)	563 (%56,3)
500	579 (%57,9)	689 (%68,9)	581 (%58,1)	539 (%53,9)
600	600 (%60,0)	628 (%62,8)	602 (%60,2)	567 (%66,7)
700	573 (%57,3)	640 (%64,0)	576 (%57,6)	532 (%53,2)
800	586 (%58,6)	601 (%60,1)	587 (%58,7)	558 (%55,8)
900	575 (%57,5)	617 (%61,7)	576 (%57,6)	539 (%53,9)
1.000	556 (%55,6)	611 (%61,1)	556 (%55,6)	529 (%52,9)
Genel	22189 (%65,1)	25701 (%75,4)	23025 (%67,5)	19712 (%57,8)

ARR: Mutlak risk azaltma.

TABLO 4: NNT ile ilgili istatistikler.

n	<i>NNT belirsiz</i>	<i>NNT < -1</i>	<i>NNT = 1</i>	<i>NNT > 1</i>
10	48 (%16,8)	119 (%41,6)	9 (%3,1)	110 (%38,5)
15	75 (%9,2)	370 (%45,4)	14 (%1,7)	356 (%43,7)
20	62 (%6,2)	471 (%47,1)	8 (%0,8)	459 (%45,9)
25	42 (%4,2)	488 (%48,8)	7 (%0,7)	463 (%46,3)
30	35 (%3,5)	461 (%46,1)	4 (%0,4)	500 (%50,0)
35	19 (%1,9)	486 (%48,6)	5 (%0,5)	490 (%49,0)
40	21 (%2,1)	490 (%49,0)	3 (%0,3)	486 (%48,6)
45	15 (%1,5)	503 (%50,3)	0 (%0,0)	482 (%48,2)
50	11 (%1,1)	489 (%48,9)	3 (%0,3)	497 (%49,7)
55	9 (%0,9)	497 (%49,7)	4 (%0,4)	490 (%49,0)
60	13 (%1,3)	493 (%49,3)	5 (%0,5)	489 (%48,9)
65	6 (%0,6)	496 (%49,6)	4 (%0,4)	494 (%49,4)
70	10 (%1,0)	480 (%48,0)	0 (%0,0)	510 (%51,0)
75	6 (%0,6)	507 (%50,7)	2 (%0,2)	485 (%48,5)
80	4 (%0,4)	485 (%48,5)	3 (%0,3)	508 (%50,8)
85	1 (%0,1)	513 (%51,3)	0 (%0,0)	486 (%48,6)
90	3 (%0,3)	511 (%51,1)	1 (%0,1)	485 (%48,5)
95	6 (%0,6)	509 (%50,9)	1 (%0,1)	484 (%48,4)
100	5 (%0,5)	465 (%46,5)	1 (%0,1)	529 (%52,9)
125	1 (%0,1)	492 (%49,2)	0 (%0,0)	507 (%50,7)
150	0 (%0,0)	500 (%50,0)	0 (%0,0)	500 (%50,0)
175	1 (%0,1)	477 (%47,7)	0 (%0,0)	522 (%52,2)
200	2 (%0,2)	513 (%51,3)	1 (%0,1)	484 (%48,4)
225	2 (%0,2)	485 (%48,5)	0 (%0,0)	513 (%51,3)
250	0 (%0,0)	480 (%48,0)	0 (%0,0)	520 (%52,0)
300	0 (%0,0)	487 (%48,7)	0 (%0,0)	513 (%51,3)
350	0 (%0,0)	486 (%48,6)	1 (%0,1)	513 (%51,3)
400	0 (%0,0)	503 (%50,3)	0 (%0,0)	497 (%49,7)
450	0 (%0,0)	523 (%52,3)	1 (%0,1)	476 (%47,6)
500	0 (%0,0)	501 (%50,1)	0 (%0,0)	499 (%49,9)
600	0 (%0,0)	527 (%52,7)	0 (%0,0)	473 (%47,3)
700	0 (%0,0)	502 (%50,2)	0 (%0,0)	498 (%49,8)
800	0 (%0,0)	522 (%52,2)	1 (%0,1)	477 (%47,7)
900	0 (%0,0)	495 (%49,5)	0 (%0,0)	505 (%50,5)
1.000	0 (%0,0)	500 (%50,0)	0 (%0,0)	500 (%50,0)
Genel	397 (%1,2)	16826 (%49,3)	78 (%0,2)	16800 (%49,3)

NNT: Tedavi için gerekli sayı.

TABLO 5: NNT için yöntemlere göre $AS > ÜS$ durum sayısı.

n	Wald	Newcombe	Agresti-Caffo	Anbar
10	130 (%45,5)	254 (%88,8)	240 (%83,9)	80 (%28,0)
15	392 (%48,1)	740 (%90,8)	619 (%61,9)	234 (%28,7)
20	495 (%49,5)	903 (%90,3)	700 (%70,0)	279 (%27,9)
25	481 (%48,1)	863 (%86,3)	649 (%64,9)	270 (%27,0)
30	481 (%48,1)	847 (%84,7)	604 (%60,4)	260 (%26,0)
35	444 (%44,4)	814 (%81,4)	558 (%55,8)	248 (%24,8)
40	450 (%45,0)	794 (%79,4)	544 (%54,4)	257 (%25,7)
45	439 (%43,9)	764 (%76,4)	503 (%50,3)	228 (%22,8)
50	405 (%40,5)	734 (%73,4)	463 (%46,3)	195 (%19,5)
55	419 (%41,9)	708 (%70,8)	469 (%46,9)	220 (%22,0)
60	404 (%40,4)	724 (%72,4)	457 (%45,7)	222 (%22,2)
65	385 (%38,5)	718 (%71,8)	426 (%42,6)	202 (%20,2)
70	390 (%39,0)	668 (%66,8)	413 (%41,3)	200 (%20,0)
75	358 (%35,8)	657 (%65,7)	397 (%39,7)	176 (%17,6)
80	369 (%36,9)	658 (%65,8)	402 (%40,2)	196 (%19,6)
85	339 (%33,9)	642 (%64,2)	370 (%37,0)	168 (%16,8)
90	341 (%34,1)	616 (%61,6)	369 (%36,9)	184 (%18,4)
95	308 (%30,8)	593 (%59,3)	344 (%34,4)	157 (%15,7)
100	332 (%33,2)	599 (%59,9)	354 (%35,4)	156 (%15,6)
125	309 (%30,9)	561 (%56,1)	337 (%33,7)	156 (%15,6)
150	258 (%25,8)	512 (%51,2)	283 (%28,3)	123 (%12,3)
175	273 (%27,3)	483 (%48,3)	282 (%28,2)	128 (%12,8)
200	272 (%27,2)	485 (%48,5)	283 (%28,3)	141 (%14,1)
225	227 (%22,7)	421 (%42,1)	233 (%23,3)	109 (%10,9)
250	222 (%22,2)	397 (%39,7)	228 (%22,8)	94 (%9,4)
300	187 (%18,7)	356 (%35,6)	190 (%19,0)	86 (%8,6)
350	172 (%17,2)	326 (%32,6)	178 (%17,8)	98 (%9,8)
400	177 (%17,7)	327 (%32,7)	185 (%18,5)	86 (%8,6)
450	182 (%18,2)	346 (%34,6)	185 (%18,5)	96 (%9,6)
500	187 (%18,7)	341 (%34,1)	191 (%19,1)	85 (%8,5)
600	156 (%15,6)	288 (%28,8)	156 (%15,6)	85 (%8,5)
700	150 (%15,0)	269 (%26,9)	156 (%15,6)	65 (%6,5)
800	134 (%13,4)	237 (%23,7)	135 (%13,5)	73 (%7,3)
900	145 (%14,5)	252 (%25,2)	148 (%14,8)	72 (%7,2)
1.000	115 (%11,5)	207 (%20,7)	115 (%11,5)	51 (%5,1)
Genel	10528 (%30,9)	19104 (%56,0)	12166 (%35,7)	5480 (%16,1)

NNT: Tedavi için gerekli sayı.

TABLO 6: Yöntemlere göre $AS \leq NNT \leq \dot{U}S$ durum sayısı.

n	Wald	Newcombe	Agresti-Caffo	Anbar
10	58 (%20,3)	5 (%1,7)	23 (%8,0)	92 (%32,2)
15	182 (%22,3)	22 (%2,7)	98 (%12,0)	275 (%33,7)
20	228 (%22,8)	39 (%3,9)	142 (%14,2)	353 (%35,3)
25	234 (%23,4)	66 (%6,6)	170 (%17,0)	346 (%34,6)
30	278 (%27,8)	65 (%6,5)	219 (%21,9)	386 (%38,6)
35	278 (%27,8)	87 (%8,7)	223 (%22,3)	380 (%38,0)
40	268 (%26,8)	89 (%8,9)	221 (%22,1)	368 (%36,8)
45	280 (%28,0)	116 (%11,6)	252 (%25,2)	373 (%37,3)
50	297 (%29,7)	128 (%12,8)	271 (%27,1)	400 (%40,0)
55	287 (%28,7)	143 (%14,3)	268 (%26,8)	384 (%38,4)
60	292 (%29,2)	143 (%14,3)	269 (%26,9)	387 (%38,7)
65	314 (%31,4)	129 (%12,9)	293 (%29,3)	396 (%39,6)
70	321 (%32,1)	148 (%14,8)	309 (%30,9)	412 (%41,2)
75	318 (%31,8)	174 (%17,4)	306 (%30,6)	404 (%40,4)
80	323 (%32,3)	168 (%16,8)	312 (%31,2)	405 (%40,5)
85	323 (%32,3)	174 (%17,4)	310 (%31,0)	404 (%40,4)
90	323 (%32,3)	193 (%19,3)	313 (%31,3)	405 (%40,5)
95	346 (%34,6)	198 (%19,8)	329 (%32,9)	412 (%41,2)
100	356 (%35,6)	190 (%19,0)	344 (%34,4)	442 (%44,2)
125	346 (%34,6)	216 (%21,6)	336 (%33,6)	422 (%42,2)
150	371 (%37,1)	247 (%24,7)	360 (%36,0)	442 (%44,2)
175	373 (%37,3)	245 (%24,5)	371 (%37,1)	451 (%45,1)
200	350 (%35,0)	275 (%27,5)	342 (%34,2)	420 (%42,0)
225	401 (%40,1)	282 (%28,2)	397 (%39,7)	464 (%46,4)
250	408 (%40,8)	287 (%28,7)	405 (%40,5)	476 (%47,6)
300	419 (%41,9)	310 (%31,0)	417 (%41,7)	471 (%47,1)
350	427 (%42,7)	329 (%32,9)	425 (%42,5)	462 (%46,2)
400	401 (%40,1)	342 (%34,2)	397 (%39,7)	451 (%45,1)
450	395 (%39,5)	345 (%34,5)	394 (%39,4)	437 (%43,7)
500	421 (%42,1)	311 (%31,1)	419 (%41,9)	461 (%46,1)
600	400 (%40,0)	372 (%37,2)	398 (%39,8)	433 (%43,3)
700	427 (%42,7)	360 (%36,0)	424 (%42,4)	468 (%46,8)
800	414 (%41,4)	399 (%39,9)	413 (%41,3)	442 (%44,2)
900	425 (%42,5)	383 (%38,3)	424 (%42,4)	461 (%46,1)
1.000	444 (%44,4)	388 (%38,8)	444 (%44,4)	471 (%47,1)
Genel	11728 (%34,4)	7368 (%21,6)	11038 (%32,4)	14256 (%41,8)

NNT: Tedavi için gerekli sayı; AS: Alt sınır; ÜS: Üst sınır.

[Tablo 5](#)'ten de görüldüğü üzere, örneklem büyüklüğü arttıkça alt sınır değerinin üst sınır değerinden büyük olduğu durum sayısı azalmaktadır. Alt sınır değeri üst sınır değerinden büyük olan güven aralığı sayısı bakımından çalışmada dikkate alınan yöntemler iyiden kötüye doğru Anbar, Wald, Agresti-Caffo ve Newcombe şeklinde sıralanmaktadır. Yöntemlere göre elde edilen güven aralıklarının NNT değerini kapsayıp kapsamaması ile ilgili çalışmadan elde edilen sonuçlar örneklem büyüklüklerine göre incelenmiş ve [Tablo 6](#)'da verilmiştir.

[Tablo 6](#)'dan da görüldüğü üzere, örneklem büyüklüğü arttıkça tüm yöntemler için NNT değerini kapsayan güven aralıklarının sayısı artmaktadır. Kapsama bakımından, çalışmada dikkate alınan yöntemler iyiden kötüye doğru Anbar, Wald, Agresti-Caffo ve Newcombe şeklinde sıralanmaktadır.

TARTIŞMA

Klinisyenlerin, farklı tedavilerin etkinliğini değerlendirebilmeleri için yaygın olarak kullanılan istatistiksel ölçütlerin anlaşılması gereklidir. NNT, belirli bir tedavinin klinik etkisinin izole edilmiş bir mutlak ölçüsünden ziyade, tek bir çalışma içindeki 2 tedavi seçeneği arasındaki tek bir karşılaştırmanın sonuçlarına özgüdür.²⁹ Risk farkı olarak da adlandırılan ARR, bir olay ile ilgili tedavi grubundaki mutlak risk ile kontrol grubundaki mutlak risk arasındaki farktır. Bir tedavi etkili ve istenmeyen bir olay riskini azaltıyorsa mutlak bir risk azalması, eğer tedavi işe yaramazsa ve aslında olayın riskini artırıyorsa mutlak bir risk artışı olacaktır. ARR'nin klinik önemini kavramsallaştırmak zor olabilir.³⁰ NNT, yeni önerilen tedavi için olumlu veya olumsuz içeriğe sahip olduğundan bu kavramı bir tedaviden diğerine değiştirilmesi gereken sayı olarak kullanmak muhtemelen daha doğru olacaktır. Değiştirilmesi gereken sayı kavramı, plasebo kontrollü denemeler düşünüldüğünde özellikle önemlidir.³¹ NNT, tedavilerin etkinliğini sıralamak için istatistiksel olarak sağlam ve kolayca yorumlanabilir bir ölçü olarak geliştirilmiş olmasına rağmen hesaplama ve yorumlamada belirli zaafiyetlere sahiptir. Bu zaafiyetler;

- Yetersiz hesaplama yöntemlerinin kullanılması, yanlış ve yanıltıcı sonuçlara yol açabilir.
- Yalnızca türetildiği çalışma geçerli olduğunda geçerlidir.
- Sonsuz bir değere sahip olabilir.
- Kesinliğini tahmin etmek zordur.
- Çarpık bir güven aralığına sahiptir.
- Plasebo grubundaki etkinlik düzeyine duyarlıdır.
- Ölçülen sonuçların seçimine bağlıdır.
- İkili sonuç için seçilen kesme değerine bağlıdır.
- Sonucun zaman noktasına bağlıdır.
- Çalışma ilacının yanı sıra çalışmanın içindeki ve dışındaki faktörlere bağlıdır.
- Klinik pratikte tedavinin etkinliğini yansıtmaz.
- Metaanalizler için problemler yaratır ve başarılı çalışmalara karşı başarısız çalışmaların orantısız bir etkisine neden olur.

şeklinde sıralanabilir.³² Çarpık bir güven aralığına sahip olma, hesaplama yöntemlerinin yetersizliği, örneklem büyüklüğünün etkisi gibi zaafiyetler bu çalışmada da belirlenmiştir. NNT zayıf özelliklere sahiptir ve en iyi ihtimalle ARR ile aynı bilgiyi içermektedir. NNT kavramı, isminden dolayı insanlara, olasılık içeriği olmayan kesin bir sayı olduğunu düşündürmektedir. Buna karşılık ARR, tahmin edicide bulunan hem olasılıksal unsurdan hem de karşılaştırmalı unsurdan açıkça bahsetmektedir. ARR, iyi istatistiksel özelliklere sahip mutlak bir ölçü olduğundan, hem istatistiksel çıkarımı hem de bilimsel sonuçları ARR'ye dayandırmak daha uygundur.³³ Temel risk azaldıkça ARR değeri küçülür, NNT değeri ise büyür. Bu nedenle nispeten küçük temel risk için potansiyel NNT değeri nispeten büyük olma eğilimindedir. ARR 0'a

yaklaştıkça NNT değeri büyür ve NNT için güven aralığı genişler. NNT değeri açısından bu durum, yeni tedavi kontrolü ile karşılaştırıldığında ne kadar az etkiliyse kontrole kıyasla yeni tedavinin daha etkili olduğu kesinliği o kadar az olur şeklinde yorumlanabilir.³⁴ NNT, randomize veya gözlemsel bir çalışmada, basit bir hesaplamanın sonucu olarak tahmin edilebilir. Ancak kontrol grubu, tedavi grubu, tedavinin etkinliği ve tedavinin herhangi bir yan etkisinin olup olmadığı gibi faktörlerin dikkate alınmaması, NNT'nin yorumlanmasında ve kullanılmasında ciddi hatalara yol açar. Gerek bir tedavinin potansiyel etkisini gerekse hastalığın halk sağlığı üzerindeki yükünü tanımlamak için NNT kavramına dayalı yeni etkinlik ölçümler ortaya konulmaktadır.⁴

SONUÇ

Çalışmadan elde edilen bulgulara göre NNT ve NNT için önerilen güven aralıkları örneklem büyüklüğünden etkilenmektedir. NNT ve NNT için önerilen güven aralığı hesaplama yöntemleri tatmin edici sonuçlar sunmamaktadır. Bu duruma, NNT ve NNT için önerilen güven aralığı hesaplama yöntemlerinin sebep olduğu, dolayısıyla hesaplama ve yorumlamada mevcut zaafiyetleri giderecek yeni yöntemlerin ortaya konulması gerektiği söylenebilir. Bütün bu olumsuzluklara rağmen alt sınır değerlerinin üst sınır değerlerinden küçük olduğu durum sayısı ve güven aralığının NNT değerini kapsamı kriterleri bakımından Anbar yönteminin çalışmada dikkate alınan diğer yöntemlere tercih edilmesi gerektiği sonucuna ulaşılmıştır.

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. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med*. 1988;318(26):1728-33. [[Crossref](#)] [[PubMed](#)]
2. Perez-Padilla R, Patino CM, Ferreira JC. Number needed to treat: a useful statistic to evaluate the impact of an intervention. *J Bras Pneumol*. 2019;45(6):e20190371. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
3. Stang A, Poole C, Bender R. Common problems related to the use of number needed to treat. *J Clin Epidemiol*. 2010;63(8):820-5. [[Crossref](#)] [[PubMed](#)]
4. Westreich D, Mooney SJ. Number (of whom?) needed to treat (with what?): exposures, population interventions, and the number needed to treat. *Epidemiology*. 2019;30 Suppl 2:S55-9. [[Crossref](#)] [[PubMed](#)]
5. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ*. 1995;310(6977):452-4. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
6. Tschudy MM, Rowe PC. Research and statistics: number needed to treat and intention to treat analysis. *Pediatr Rev*. 2010;31(9):380-2. [[Crossref](#)] [[PubMed](#)]
7. Elliott MH, Skydel JJ, Dhruva SS, Ross JS, Wallach JD. Characteristics and reporting of number needed to treat, number needed to harm, and absolute risk reduction in controlled clinical trials, 2001-2019. *JAMA Intern Med*. 2021;181(2):282-4. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
8. Bowry SK, Schoder V, Apel C. An inadvertent but explicable error in calculating number needed to treat for reporting survival data. *J Am Soc Nephrol*. 2014;25(5):875-6. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
9. Nguyen C, Naunton M, Thomas J, Todd L, McEwen J, Bushell M. Availability and use of number needed to treat (NNT) based decision aids for pharmaceutical interventions. *Explor Res Clin Soc Pharm*. 2021;2:100039. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
10. Biswas A. Number needed to treat. *J Pract Cardiovasc Sci*. 2017;3(2):106-8. [[Crossref](#)]
11. Altman DG. Confidence intervals for the number needed to treat. *BMJ*. 1998;317(7168):1309-12. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
12. Veroniki AA, Bender R, Glasziou P, Straus SE, Tricco AC. The number needed to treat in pairwise and network meta-analysis and its graphical representation. *J Clin Epidemiol*. 2019;111:11-22. [[Crossref](#)] [[PubMed](#)]
13. Saver JL, Lewis RJ. Number needed to treat: conveying the likelihood of a therapeutic effect. *JAMA*. 2019;321(8):798-9. [[Crossref](#)] [[PubMed](#)]
14. Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. *Biol Psychiatry*. 2006;59(11):990-6. [[Crossref](#)] [[PubMed](#)]
15. Furukawa TA, Leucht S. How to obtain NNT from Cohen's d: comparison of two methods. *PLoS One*. 2011;6(4):e19070. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
16. Vancak V, Goldberg Y, Levine SZ. Systematic analysis of the number needed to treat. *Stat Methods Med Res*. 2020;29(9):2393-410. [[Crossref](#)] [[PubMed](#)]
17. Wald NJ, Morris JK. Two under-recognized limitations of number needed to treat. *Int J Epidemiol*. 2020;49(2):359-60. [[Crossref](#)] [[PubMed](#)]
18. Weeks DL, Noteboom JT. Using the number needed to treat in clinical practice. *Arch Phys Med Rehabil*. 2004;85(10):1729-31. [[Crossref](#)] [[PubMed](#)]
19. Bender R. Calculating confidence intervals for the number needed to treat. *Control Clin Trials*. 2001;22(2):102-10. [[Crossref](#)] [[PubMed](#)]
20. Ludwig L, Darmon P, Guerci B. Computing and interpreting the Number Needed to Treat for Cardiovascular Outcomes Trials: perspective on GLP-1 RA and SGLT-2i therapies. *Cardiovasc Diabetol*. 2020;19(1):65. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
21. Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat Med*. 1998;17(8):873-90. [[Crossref](#)] [[PubMed](#)]
22. 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](#)]
23. Anbar D. On estimating the difference between two probabilities, with special reference to clinical trials. *Biometrics*. 1983;39(1):257-62. [[Crossref](#)] [[PubMed](#)]
24. Agresti A, Caffo B. Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *Am Stat*. 2000;54(4):280-8. [[Crossref](#)]
25. Ebrahim S, Smith GD. The 'number need to treat': does it help clinical decision making? *J Hum Hypertens*. 1999;13(11):721-4. [[Crossref](#)] [[PubMed](#)]
26. Barratt A, Wyer PC, Hatala R, McGinn T, Dans AL, Keitz S, et al; Evidence-Based Medicine Teaching Tips Working Group. Tips for learners of evidence-based medicine: 1. Relative risk reduction, absolute risk reduction and number needed to treat. *CMAJ*. 2004;171(4):353-8. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
27. San Laureano JA, Briganti EM, Colville DJ. Number needed to treat: a useful new method of assessing the magnitude of treatment effect and its application to the management of diabetic retinopathy. *Aust N Z J Ophthalmol*. 1999;27(2):137-42. [[Crossref](#)] [[PubMed](#)]
28. Barrowman NJ. Missing the point (estimate)? Confidence intervals for the number needed to treat. *CMAJ*. 2002;166(13):1676-7. [[PubMed](#)] [[PMC](#)]
29. Mendes D, Alves C, Batel-Marques F. Number needed to treat (NNT) in clinical literature: an appraisal. *BMC Med*. 2017;15(1):112. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
30. Scott I. Interpreting risks and ratios in therapy trials. *Aust Prescr*. 2008;31(1):12-6. [[Crossref](#)]

31. de Craen AJ, Vickers AJ, Tijssen JG, Kleijnen J. Number-needed-to-treat and placebo-controlled trials. *Lancet*. 1998;351(9099):310. [\[Crossref\]](#) [\[PubMed\]](#)
32. Katz N, Paillard FC, Van Inwegen R. A review of the use of the number needed to treat to evaluate the efficacy of analgesics. *J Pain*. 2015;16(2):116-23. [\[Crossref\]](#) [\[PubMed\]](#)
33. Hutton JL. Number needed to treat: properties and problems. *J R Stat Soc Ser A Stat Soc*. 2000;163(3):381-402. [\[Crossref\]](#)
34. Raju GK, Khozin S, Gurumurthi K, Domike R, Woodcock J. Patient-centered approach to benefit-risk characterization using number needed to benefit and number needed to harm: advanced non-small-cell lung cancer. *JCO Clin Cancer Inform*. 2020;4:769-83. [\[Crossref\]](#) [\[PubMed\]](#) [\[PMC\]](#)