



## RESEARCH ARTICLE

### Klinik çalışmalarda yapay zekâ uygulamalarının raporlanmasında CONSORT-AI protokolünün kullanımı

Mustafa Agah Tekindal<sup>1</sup>, Melike Tekindal<sup>2</sup>, Mert Demirsöz<sup>3</sup>, Zeynep Özel<sup>3</sup>

<sup>1</sup>İzmir Kâtip Çelebi Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı, İzmir, Türkiye  
<sup>2</sup>İzmir Kâtip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi Sosyal Hizmet Anabilim Dalı, İzmir, Türkiye  
<sup>3</sup>Selçuk Üniversitesi Veteriner Fakültesi Biyoistatistik Anabilim Dalı, 42130, Konya, Türkiye

Geliş: 20.09.2021, Kabul: 04.02.2022  
\*mustafaagah.tekindal@ikc.edu.tr

### Use of the CONSORT-AI protocol in reporting artificial intelligence applications in clinical trials

Eurasian J Vet Sci, 2022, 38, 2, 66-74  
DOI: 10.15312/EurasianJVetSci.2022.366

#### Öz

**Amaç:** Yapay zekâ 21. Yüzyılda teknolojinin gelişimi ile tıpta dahil olmak üzere birçok bilim dalında kullanılmaya başlanmıştır. Yapılan araştırma sonuçlarının uluslararası standartlara uygun nitelikte raporlanması, söz konusu araştırmaları anlama, yorumlama ve eleştirel olarak değerlendirme sürecine anlamlı katkılar sağlayacaktır. CONSORT-AI (Consolidated Standards of Reporting Trials- Artificial Intelligence [Araştırma Raporlamada Konsolide Standartlar- Yapay Zekâ]) sayısı hızla artan yapay zekâ uygulamalarına bağlı olarak tıp ve sağlık alanında kullanımı yaygınlaşan bir raporlama protokolüdür. Bu çalışmada Türkçe uyarlaması yapılan CONSORT-AI protokolünün araştırmacılar tarafından doğru bir şekilde anlaşılıp anlaşılmadığını değerlendirmek amaçlanmıştır.

**Gereç ve Yöntem:** TR Dizin dergilerden seçilen 22 makale CONSORT-AI protokolü ile değerlendirici konumunda bulunan yazarlar tarafından değerlendirilmiş ve değerlendirme sonuçları yazarlar arası uyum analizi ile test edilmiştir.

**Bulgular:** Elde edilen bulgular, değerlendiriciler arası uyumun istatistiksel olarak anlamlı olduğunu göstermiştir. Bununla birlikte incelenen araştırmalarda CONSORT-AI protokolündeki ilkeleri tümüyle gözetilen bir makaleye de rastlanamamıştır.

**Öneri:** Çalışmanın yapay zekâ odaklı araştırma çıktılarında izlenmesi gereken ilkelere dikkat çekerek tıp ve sağlık alanında görev yapan bilgi profesyonellerine, konuyla ilgili araştırmacılara ve araştırmaları değerlendirme konumunda bulunan editör ve hakemlere katkı sağlaması beklenmektedir.

**Anahtar kelimeler:** Yapay zekâ, klinik çalışmalar, protokol, CONSORT-AI.

#### Abstract

**Aim:** Artificial intelligence has begun to be used in many branches of science, including medicine, with the development of technology in the 21st century. However, these studies are reported outside of a certain standard and this situation causes the studies to have a complex structure. Reporting the research results in accordance with international standards will make meaningful contributions to the process of understanding, interpreting and critically evaluating the research in question. CONSORT-AI (Consolidated Standards of Reporting Trials- Artificial Intelligence) is a reporting protocol that has become widespread in medicine and health due to the rapidly increasing number of artificial intelligence applications. In this study, it was aimed to evaluate whether the CONSORT-AI protocol, which was adapted into Turkish, was correctly understood by the researchers.

**Materials and Methods:** 22 articles selected from TR Index journals were evaluated by the authors who were in the evaluator position with the CONSORT-AI protocol, and the evaluation results were tested with inter-author agreement analysis.

**Results:** The obtained findings showed that the agreement between raters was statistically significant. However, in the studies examined, no article was found that fully observes the principles in the CONSORT-AI protocol.

**Conclusion:** It is expected that the study will contribute to information professionals working in the field of medicine and health, researchers related to the subject, and editors and referees who are in the position of evaluating research by drawing attention to the principles to be followed in artificial intelligence-oriented research outputs.

**Keywords:** Artificial intelligence, clinical studies, protocol, CONSORT-AI.



## Giriş

Tıp bilgisinin ikiye katlanma süresinin 1950'de 50 yıl, 1980'de 7 yıl, 2010'da 3,5 yıl, 2020'de ise 0,2 yıl yani sadece 73 gün olduğu öne sürülmektedir (Densen 2011). Buna karşılık pratisyen bir hekimin güncel literatürü takip edebilmesi için günde ortalama 19 makale okuması gerekmektedir. Yoğun iş yükleri göz önünde bulundurulduğunda, belirtilen sayıda makale okumanın hekimler açısından oldukça güç olduğu söylenebilir (Masic ve ark 2008).

Yeni tedaviler, tanılama yöntemleri ve prosedürler gibi sürekli gelişmeler, tıp ve sağlık alanında doğru ve kapsamlı bilgilere zamanında erişimi daha anlamlı hale gerektirmektedir. Başarılı bir sağlık hizmeti, başarılı sağlık araştırmalarına; söz konusu araştırmalar ise etkin şekilde yönetilen bilgi kaynaklarına ve bu kaynakların kapsadığı bilimsel bilgilere bağlıdır. Tıp Kütüphanesi Derneği (Medical Library Association-MLA 2021), tıp ve sağlık bilimleri alanında görev yapan bilgi profesyonellerinin 21. Yüzyılda taşıması gereken niteliklere ve mesleki işlevlerine açıklık getirirken, hasta bakımı, eğitim, araştırma ve yayın faaliyetlerini bilgi hizmetleriyle desteklemenin yanı sıra, sağlık çalışanlarına bilgiye nasıl erişeceklerini ve değerlendireceklerini öğretme; literatürü ve dolayısıyla araştırma çıktılarını eleştirel olarak değerlendirme rollerine de dikkat çeker. Dahası yapılan araştırmaların niteliğini artırıcı stratejilerin ve araçların geliştirilmesi, literatür taraması, eleştirel değerlendirme, kanıtları entegre etme gibi kanıta dayalı tıp uygulamaları, hekim-bilgi profesyoneli iş birliğini gerektiren başlıca etkinlikler arasındadır.

Klinisyen hekimler bilimsel makaleleri kullanarak özeldir hastayı, genelde ulusal ve uluslararası sağlık hizmetlerini iyileştirmektedir. Bu nedenle, bilimsel makalelerin hekimlerin sundukları hizmetleri daha verimli bir şekilde destekleyebilecek yeni teknik ve süreçleri benimsemesi giderek daha anlamlı hale gelmektedir. İyi kurgulanmış araştırma çıktıları, okuyucuların çalışmanın tam olarak nasıl yapıldığını, araştırma tasarımını, araştırmaların yürütülmesini ve çıktıların yayınlanma sürecini anlamasını sağlar. Bu durum eksiksiz raporlama ölçütlerini temel alan araştırmaların ve doğru çıktılar üretmeyi sağlayan rehberlerin değerini artırmaktadır. Bu araştırmada, son yıllarda sayısı hızla artan tıpta yapay zekâ uygulamalarını konu edinen makaleler, klinik araştırma çıktıları için şeffaflığı ve eksiksizliği teşvik eden ve okuyuculara klinik araştırma raporlarını anlama, yorumlama ve eleştirel olarak değerlendirme konusunda yardımcı olmayı amaçlayan CONSORT-AI (Consolidated Standards of Reporting Trials- Artificial Intelligence [Araştırma Raporlamada Konsolide Standartlar- Yapay Zekâ]) protokolü araştırmacılara çalışma ile ilgili yön gösteren bir protokoldür. Bu protokolün Türkçe literatürde de kullanımını yaygınlaştırma hedefinden yola çıkarak araştırmada Türkçe uyarlamasının yapılarak değerlendirilmeleri tarafından doğru anlaşılıp anlaşılmadığını kontrol etmek amaçlanmaktadır.

## Gereç ve Yöntem

### CONSORT-AI

İnsan beyninin çalışma özelliğini modelleyerek insanın zekâsını bilgisayar aracılığı ile taklit eden yapay zekâ uygulamaları 1969 yılından bu yana tıp alanında çeşitli amaçlarla kullanılmaktadır (Ahuja 2019, Serhatlıoğlu ve Hardalac 2009). Yapay zekâ uygulamaları tıbbi girişimlerde pratiklik sağlamanın yanı sıra miktarı her geçen gün artan hasta verilerini toplama, analiz etme ve bu verileri hastanın tıbbi gereksinimlerinin anlaşılmasında kullanabilme yeteneği de sunmaktadır. Hekimlerin zaman tüketimini konu edinen bir araştırma (Sinsky ve ark 2016) zamanın büyük bir kısmını elektronik hastane kayıtları ve masa başı çalışmalarına harcadıklarını öne sürmektedir. Bu durum, yapay zekâ uygulamalarıyla bütünleşen sağlık hizmetlerinin hekimlerin zamanını daha üretken kullanmasına ve dolayısıyla daha etkin sağlık hizmetlerinin ortaya çıkmasına katkı sağlayacağını düşündürmektedir (Amisha ve ark 2019).

Tıpta yapay zekâ uygulamalarını konu edinen araştırmaları standart bir yapıda raporlaştırmak, okuyucular ve araştırmacılar için eleştirel değerlendirme ve yorumlama gibi süreçlerin daha etkili gerçekleşmesine neden olurken, araştırma verimliliğini de artırmaktadır. Araştırma çıktıların üretilmesi ve sunumu, günümüzde protokoller aracılığıyla eksiksiz ve şeffaf bir şekilde gerçekleştirilmekte. Bu bağlamda öncül girişimlerden CONSORT, birbirinden farklı araştırma çıktıların iyileştirilmesini ve sağlık hizmetlerinde karar verme aşamasında kullanılan araştırma verilerinin kalitesini artırmayı hedefleyen bir çabanın ürünüdür (CONSORT 2021). CONSORT, 25 maddelik bir protokol ve bir akış diyagramından oluşur. CONSORT'un ana ürünü, randomize (rastgele) çalışmaların raporlanması için kanıta dayalı, minimum bir öneri seti olan CONSORT protokoldür. Yazarlara klinik deneme bulgularına dayalı rapor hazırlamaları için standart bir yol sunarak, eksiksiz ve şeffaf raporlamayı kolaylaştırır. Ayrıca CONSORT, kanıta dayalı tıp uygulamalarının temel beklentilerinden eleştirel değerlendirme ve yorumlamayı da sağlayarak hem araştırmacılara hem de araştırmaları değerlendirenlere yardımcı olur. Protokol maddeleri, klinik araştırmanın nasıl tasarlandığını, nasıl analiz edildiğini ve nasıl yorumlandığını raporlamaya odaklanırken; akış diyagramı, klinik deneme boyunca tüm katılımcıların nasıl ilerlediklerini gösterir (CONSORT 2021).

CONSORT'un CONSORT-AI adıyla anılan yapay zekâ uzantısı, klinik müdahaleleri bir yapay zekâ bileşeniyle değerlendiren araştırmalar için kullanılan yeni bir raporlama kılavuzudur. CONSORT-AI bilinen CONSORT maddelerine ek olarak 14 yeni maddeyi içermektedir. CONSORT-AI, araştırmacılar tarafından yapay zekâ müdahalesinin bütünleştirildiği ortam, yapay zekâ uygulamasına bağlı girdilerin ve çıktıların işlenmesi, hata analizi, gerekli talimatlar ve beceriler, insan-yapay

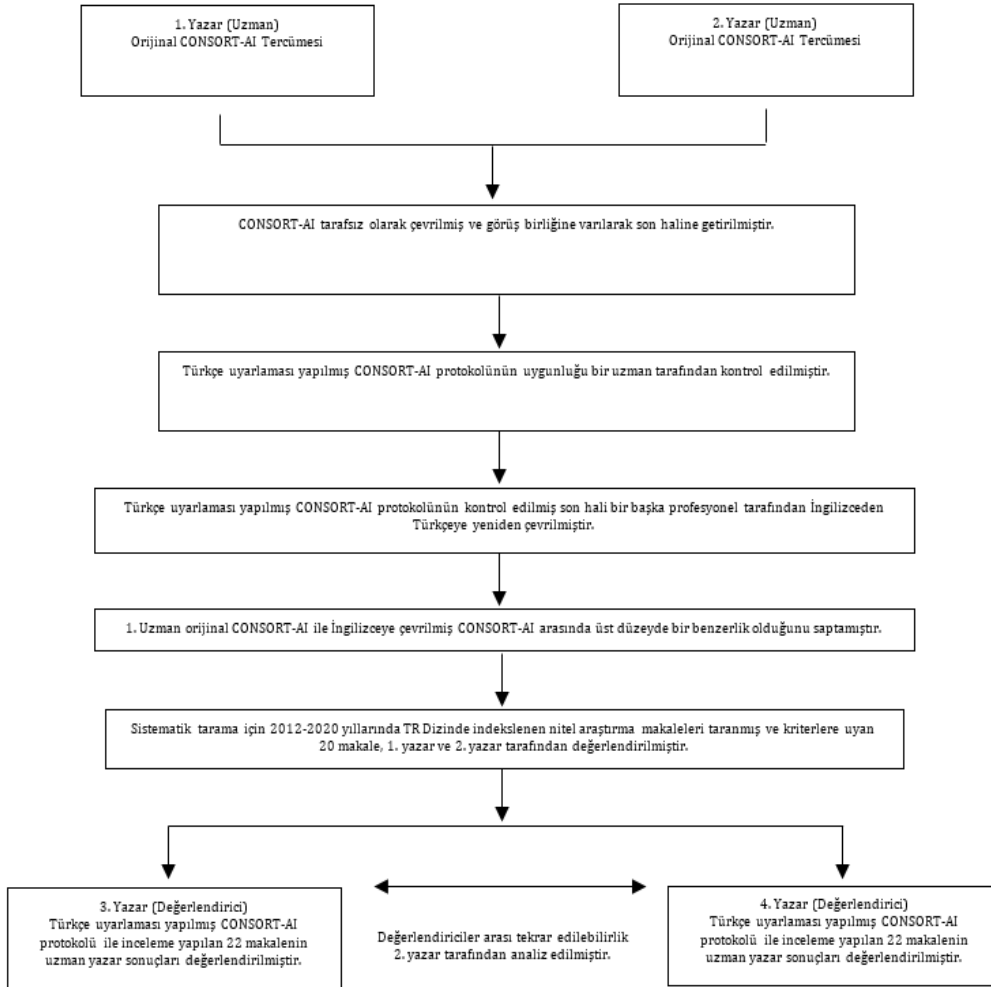
zekâ etkileşimi gibi yapay zekâ kullanılan çalışmaların açıkça ve anlaşılır şekilde tanımlanmasını önerir (Liu ve ark 2020). Bu özellikleriyle CONSORT-AI, yapay zekâ ile bütünleşen klinik müdahalelerin raporlanmasında bütünlüğü ve eksiksizliği teşvik etmektedir. Editörlerin ve hakemlerin yanı sıra genel okuyucuların da klinik araştırma tasarımının kalitesini anlamasına, yorumlamasına ve eleştirel bir şekilde değerlendirmesine yardımcı olur (İbrahim ve ark 2021).

Değerlendirmeler kapsamında CONSORT-AI protokolü okuyucu tarafından daha anlaşılır çalışmalar oluşmasına zemin hazırlamaktadır. Bu doğrultuda tıp alanında TR Dizin dergilerden seçilen 22 makale CONSORT-AI protokolü ile iki araştırmacı tarafından protokole uygun şekilde değerlendirilmiştir. Değerlendirmelerde çalışmanın niteliği değil CONSORT-AI protokolünün çalışmacılar için benzer çıkarımlar sağlayıp sağlamadığı incelenmiştir. Bu değerlendirmeler sonucunda yazarların uyguladığı protokol sonuçları farklı 2 yazar tarafından da kontrol edilerek sonuçlar yansız bir şekilde elde edilmiştir. Bu değerlendirmeler çalışmanın teorik arka planını oluşturmuştur.

### Araştırmanın yöntemi

2020 yılında yayınlanan orijinal CONSORT-AI protokolü, makalenin yazarları tarafından ayrı ayrı Türkçe 'ye çevrilmiştir. Yazarlar arası görüş alışverişi ile protokolün Türkçe çevirisi gözden geçirilmiş ve çeviri metni üzerinde görüş birliğine varılmıştır. İzleyen aşamada, CONSORT-AI Türkçe protokolü, dil bilimi profesyoneli tarafından da kontrol edilerek dil birliğine ulaşılmıştır. Son olarak Türkçe uyarlaması yapılan CONSORT-AI, ters çevirme yöntemi kullanılarak bir başka dil bilimi profesyoneli tarafından Türkçe 'den İngilizce 'ye çevrilmiş ve orijinal CONSORT-AI protokolü ile karşılaştırılarak dil benzerliği netleştirilmiştir. Yapay Zekâ araştırmalarının raporlanmasında kullanılan CONSORT-AI protokolünün Türkçe uyarlaması Tablo 1'de verilmiştir. Bu süreçte çeviri ve protokolün uygulama aşamasında değerlendirmenin nasıl yapıldığı akış diyagramı Şekil 1'de özetlenmiştir.

Şekil 1 de makaleler arasındaki uyum değil Türkçe CONSORT-AI protokolünün uzmanlar arasındaki karar vermelerdeki uyumuna ilişkin bir akış diyagramıdır. Bu akış sonucunda



Şekil 1. Tıpta yapay zekâ uygulamalarına bağlı araştırma çıktılarının konsolide raporlamasında kullanılan CONSORT-AI protokolünün Türkçe uyarlamasında izlenen akış diyagramı

Tablo 1. Tıpta yapay zekâ uygulamalarına bağlı araştırma çıktılarının raporlamasında kullanılan CONSORT-AI protokolünün türkçe uyarlaması (Liu ve ark 2020)

Bölüm/Başlık	Madde No	CONSORT		CONSORT-AI	
		CONSORT 2010 Maddesi <sup>a</sup>	Madde No	CONSORT-AI Maddesi	
Başlık ve Özet					
Başlık ve Özet	1a	Başlıkta, randomize (rastgele) bir çalışma olarak tanımlanması			(i) Başlıkta ve/veya özetle müdahalenin yapay zekâ (AI) ve/veya makine öğrenimi (ML) içerdiğinin belirtilmesi ve modelin türünün belirtilmesi
	1b	Çalışma tasarımı, yöntemleri, bulguları ve sonuçlarının yapılandırılmış özeti (Detaylı bilgi için CONSORT özetler bölümünü inceleyiniz (Hopewell ve ark 2008)).		CONSORT-AI 1a,b Detaylandırma	(ii) Araştırmada yapay zekâ (AI) müdahalesinin amaçlanan kullanımını başlıkta ve/veya bir özetle belirtilmesi.
Giriş					
Arka plan ve amaçlar	2a	Bilimsel arka plan ve gerekçenin açıklanması		CONSORT-AI 2a (i) Ek	Yapay zekâ (AI) müdahalesinin amaçlanan kullanımını, amacı ve hedeflenen kullanıcıları (örneğin, sağlık uzmanları, hastalar, halk) dahil olmak üzere klinik yol bağlanımının açıklanması
	2b	Özgün amaç veya hipotezlerin belirtilmesi			
Yöntemler					
Çalışma tasarımı	3a	Ayrırma (dağıtım) oranları da dahil edilerek çalışma tasarımının tanımlanması (paralel, faktöriyel gibi)			
	3b	Çalışma başladıktan sonra yöntemlerdeki önemli değişikliklerinin nedenleri ile belirtilmesi (örneğin uygunluk kriteri gibi).			
Katılımcılar	4a	Katılımcılar için uygunluk kriterlerinin belirlenmesi		CONSORT-AI 4a (i) Detaylandırma CONSORT-AI 4a (ii) Ek	Katılımcılar düzeyinde dâhil etme ve hariç tutma kriterlerinin belirtilmesi Girdi verileri düzeyinde dâhil etme ve hariç tutma kriterlerinin belirtilmesi
	4b	Verilerin toplandığı ortamlar ve yerlerin belirtilmesi		CONSORT-AI 4b Ek	Yapay zekâ (AI) müdahalesinin çalışma ortamına nasıl entegre edildiğinin açıklanması; (ortam [saha] içi veya ortam dışı gereksinimler dahil olmak üzere)
Müdahaleler	5	Her bir grup için, müdahalelerin ne zaman ve nasıl uygulandıkları yeterli ayrıntıya sahip olacak şekilde açıklanarak özetlenmesi		CONSORT-AI 5 (i) Ek	Yapay zekâ (AI) algoritması için hangi sürümün kullanıldığının belirtilmesi
				CONSORT-AI 5 (ii) Ek	Yapay zekâ (AI) müdahalesi için girdi verilerinin nasıl elde edildiğinin ve seçildiğinin açıklanması
				CONSORT-AI 5 (iii) Ek	Düşük kaliteli veya kullanılmayan girdi verilerinin nasıl değerlendirildiğinin ve işlendiğinin açıklanması
				CONSORT-AI 5 (iv) Ek	Girdi verilerinin işlenmesinde insan-yapay zekâ (AI) etkileşimi olup olmadığının ve kullanıcıların hangi düzeyde uzmanlığa ihtiyaç duyduklarının belirtilmesi
				CONSORT-AI 5 (v) Ek	Yapay zekâ (AI) müdahalesinin çıktısının belirtilmesi
				CONSORT-AI 5 (vi) Ek	Yapay zekâ (AI) müdahalesinin çıktılarının karar almaya veya klinik uygulamanın diğer unsurlarına nasıl katkıda bulunduğunun açıklanması
Sonuçlar	6a	Önceden belirlenmiş ve tamamen tanımlanmış birincil ve ikincil sonuç ölçümleri (ölçüleri) ne zaman ve nasıl değerlendirildiği de dahil olmak üzere belirtilmesi			
	6b	Çalışmaya başladıktan sonra sonuçlarında herhangi bir değişiklik var ise nedenleri ile belirtilmesi			
Örneklem büyüklüğü	7a	Örneklem büyüklüğü nasıl belirlendiğinin belirtilmesi			
	7b	Uygulanabilir olduğunda, ara [interim] analizlerin ve durdurma kriterlerinin açıklanması			
Randomizasyon					
Dizi oluşturma	8a	Rastgele ayırma dizisi [sırası] oluşturmak için kullanılan yöntemin belirtilmesi			
	8b	Randomizasyon tipi; herhangi bir kısıtlamanın ayrıntılarının belirtilmesi (bloklama ve blok boyutu gibi)			
Dağıtım gizleme mekanizması	9	Müdahalelerde diziye gizlemek için atılan adımları açıklayan rastgele ayırma dizisini (Örneğin sıralı olarak numaraların atanması gibi) uygulamak için kullanılan mekanizmanın belirlenmesi			
Uygulama	10	Rastgele ayırma dizisini kim oluşturduğunu, katılımcıları kim kaydettiğini ve müdahaleler için katılımcıları kimin belirlediğinin (atadığının) belirtilmesi			
Körleme	11a	Eğer yapıldıysa, müdahalelere atandıktan sonra kimlerin çalışmadan çıkarıldığının (örneğin, katılımcılar, bakım sağlayanlar, sonuçları değerlendirenler) ve nasıl çıkarıldığının belirtilmesi			
	11b	Eğer müdahaleler arasında bir uyum var ise, müdahalelerin benzerliğinin tanımlanması			
İstatistiksel yöntemler	12a	Birincil ve ikincil sonuçlar için, grupların karşılaştırılmasında kullanılan istatistiksel yöntemlerin belirtilmesi			
	12b	Alt grup analizleri ve düzeltilmiş analizler gibi ek analizler için yöntemlerin belirtilmesi			



Tablo 1. Devamı

		Sonuçlar	
Katılımcı akışı (Bir diyagram şiddetle önerilir.)	13a	Her bir grup için, rastgele ayrılan, planlanan tedaviyi alan ve birincil sonuçlar için analiz edilen katılımcı sayısının belirtilmesi	
	13b	Her bir grup için, randomizasyon yapıldıktan sonra deney dışı bırakmaların deney dışı bırakılma nedenleri ile belirtilmesi	
Çalışmaya alım	14a	Çalışmaya alım ve izlem sürelerini tanımlayan tarihlerin belirtilmesi	
	14b	Çalışmanın neden sonlandırıldığının veya durdurulduğunun belirtilmesi	
Temel veriler	15	Her bir grubun temel demografik ve klinik özelliklerinin bir tablo ile gösterilmesi	
Analiz edilen sayılar	16	Her grup için, her analize dahil edilen katılımcı sayısı ve analizin orijinal atanmış (ayrılan) gruplara göre olup olmadığının belirtilmesi	
Sonuçlar ve tahmin	17a	Birincil ve ikincil sonuçlarda her grup için sonuçların, tahmini etki boyutunun ve kesinliğin (duyarlılığın) (%95 güven aralığı gibi) belirtilmesi	
	17b	İkili sonuçlar için hem mutlak hem de nispi etki boyutunun açıklanması	
Yardımcı analizler	18	Yapılan herhangi diğer analiz sonuçlarının (alt grup analizleri ve düzeltilmiş analizler dâhil olmak üzere) önceden belirlenmiş olanları açıklayıcı (planlanmamış) olanlardan ayrıracak şekilde belirtilmesi	
Zararlar	19	Her gruptaki tüm önemli olumsuzluklar veya istenmeyen etkiler (Detaylı bilgi için CONSORT zararlar bölümünü inceleyiniz (Ioannidis ve ark 2004) belirtilmesi.	CONSORT-AI 19 Ek
Tartışma			Herhangi bir performans hatası analizinin sonuçlarını ve uygun olduğunda hataların nasıl tanımlandığının açıklanması. Böyle bir analiz planlanmadıysa veya yapılmadıysa, neden yapılmadığının açıklanması.
Kısıtlılıklar	20	Çalışma kısıtlılıklarının (potansiyel önyargı kaynaklarını, belirsizlikleri [hataları] ve diğer ilgili ise analizlerin çeşitliliğini dikkate alacak şekilde) belirtilmesi.	
Genelleştirilebilirlik	21	Çalışma bulgularının genelleştirilebilirliği (dış geçerlik, uygulanabilirlik) belirtilmesi	
Yorum	22	Sonuçlarla tutarlı yorumlanması, fayda ve zararların dengelenmesi ve diğer ilgili kanıtları göz önünde bulundurulması	
Diğer Bilgiler			
Kayıt	23	Çalışma kaydının ismi ve kayıt numarasını belirtin.	
Protokol	24	Eğer mevcutsa, tam çalışma protokolüne nereden erişilebileceğini belirtin	
Finansman	25	Finansman kaynakların ve diğer desteklerin (ilaçların tedarigi gibi) rollerinin belirtilmesi	CONSORT-AI 25 Ek
			Yapay zekâ (AI) müdahalesine ve/veya koduna erişilip erişilemeyeceğini ve nasıl erişilebileceğini erişim veya yeniden kullanıma ilgili tüm kısıtlamalar dâhil olmak üzere belirtilmesi

elde edilen sonuçlar Bland-Altman yöntemi ile uzmanlar arasındaki uyum ve Sınıf içi Korelasyon Katsayısı (SKK) ile uzman değerlendirmeleri arasındaki tutarlılık değerlendirilmiştir. Bu sonuçlar teorik arka plana uygun olarak CONSORT-AI protokolünün Türkçe versiyonunun uzmanlar tarafından makale inceleme ve makale yazmada daha hızlı ve anlaşılır olabilmelerini sağlamayı planlanmaktadır.

Tablo 1'de 25 maddelik CONSORT protokolü ve daha sonra yapay zekâ uygulamaları için CONSORT maddelerine eklenen 14 yeni maddeyi içeren protokolün Türkçe versiyonuna yer verilmiştir. Bu tablo klinik çalışmalarda yapay zekâ uygulamalarında takip edilmesi gereken adımları içermektedir.

Araştırmanın ikinci aşamasında, Türkçeye uyarlanmış CONSORT-AI protokolünün doğru olarak anlaşılıp anlaşılmadığını belirlemek için seçilen 22 makale iki yazar tarafından incelenmiş ve bu inceleme sonucundan yanlılığı ortadan kaldırmak amacı ile iki yazar tarafından inceleme yapılan protokoller değerlendirilmiştir

### Araştırma verilerinin toplanması

17 Mart 2021 tarihinde yapılan tarama sonucu elde edilen ve iki araştırmacı tarafından değerlendirilen yayınlar, TR Dizin'de 2012-2020 yıllarında sağlık alanında yayınlanan ve dili Türkçe olan çalışmalar arasından seçilmiştir. Yayın taramasında yararlanılan önemli bir diğer kriter de başlık, özet ya da anahtar kelimelerde "yapay zekâ" ve/veya "sağlıkta yapay zekâ" ve/veya "makine öğrenmesi" kelimelerinin yer alması şeklinde belirlenmiştir. Bu tarama kriterlerini sağlayan toplam makale sayısı ilk etapta 112 olarak bulunmuş daha sonra 2. yazar tarafından makalelerin tümü değerlendirilmiş ve sağlıkta yapay zekâ çalışması olmadığı tespit edilen 90 makale elenmiştir. Eleme sonucunda 22 makale belirlenmiştir. Seçilen 22 makale çift kör olacak şekilde CONSORT-AI protokolünün Türkçeye çevrilmiş versiyonu yardımı ile incelenmiştir. Bu incelemenin temel amacı uzmanlar tarafından maddelerin benzer şekilde algılanıp algılanmadığını belirlemek için aralarındaki uyumun ölçülmesidir. Çalışmada Masic ve diğerlerinin (2008) belirttiği gibi potansiyel bir hekimin literatüre hâkim olması için günlük 19 makale okuması gerekliliğinden yola çıkarak 22 makalenin değerlendirilmesi yeterli bulunmuştur.



### Araştırma verilerinin analizi

Araştırma verileri analiz edilirken, benzer amaçlı yayınlardan Balak ve diğerlerinin (2018) makale değerlendirme aşaması temel alınmış bu bağlamda her bir maddeye 0 veya 1 değerleri verilerek puanlandırma yapılmıştır. Protokolde yer alan maddenin değerlendirilen makalede karşılık bulması durumunda 1 puan; karşılık bulmaması halinde ise 0 puan verilmiştir. Bu yolla elde edilen toplam puanlar değerlendirilen makaleler için ulaşılan toplam madde sayısını da içermektedir. CONSORT-AI odağında yapılan değerlendirmelere ilişkin, değerlendirici konumunda bulunan yazarlar arası uyumu ölçerken grafiksel bir yöntem olan Bland-Altman kullanılmıştır. Bu yolla, değerlendiricilerin makalelere verdiği puanlar arasındaki farkın ortalama ve standart sapma gibi değerleri çerçevesinde CONSORT-AI protokolü incelenmiştir (Schmidt ve Steindorf 2006). Değerlendiriciler arası uyumun belirlenebilmesi için Bland-Altman uyum grafiği çizilmiştir. CONSORT-AI protokolüne göre çizilen grafikte, makalelere verilen puanların farkları ve her makaleye verilen puanların ortalaması yer almaktadır. "Bland-Altman yönteminde iki teknik arasındaki fark, iki yöntemin ortalamalarına göre %95 güven sınırlarında belirlenen sapmalar çizilerek karşılaştırılır" (Genç ve ark 2012, Balak ve ark 2018, Tekindal ve ark 2015)

Değerlendirme sonunda makalelere verilen en yüksek puan 20 olmuştur. Daha sonra makalelere verilen puanların ortalaması ve standart sapması hesaplanmıştır. Makalelerin raporlanması sürecinin niteliğine ilişkin incelemenin yazarlar (1. ve 2. yazar) arasında hangi ölçüde benzer olduğunu saptamak için sınıf içi korelasyon katsayısı (SKK) kullanılmıştır. Bu katsayının 0.75'in üzerinde olması istenen bir durum olmakla birlikte ilgili değer 0 ile +1 aralığında değişir (Balak ve ark 2018, Lee ve ark 1989, McGraw ve Wong 1996). SKK belirlenirken iki yönlü rastgele etkili model ve mutlak uyum seçeneği kullanılmıştır (Genç ve ark 2012, Ateş ve ark 2009, Balak ve ark 2018).

### Bulgular

Tıpta Yapay Zekâ Uygulamalarının raporlarının yazımı için geliştirilen, Türkçeye uyarlanmış CONSORT-AI protokolüne ilişkin SKK değerleri Tablo 2'de verilmiştir. Tüm makale incelemelerinde uzmanlar arası uyum istatistik olarak anlamlı bulunmuştur ( $p < 0,05$ ).

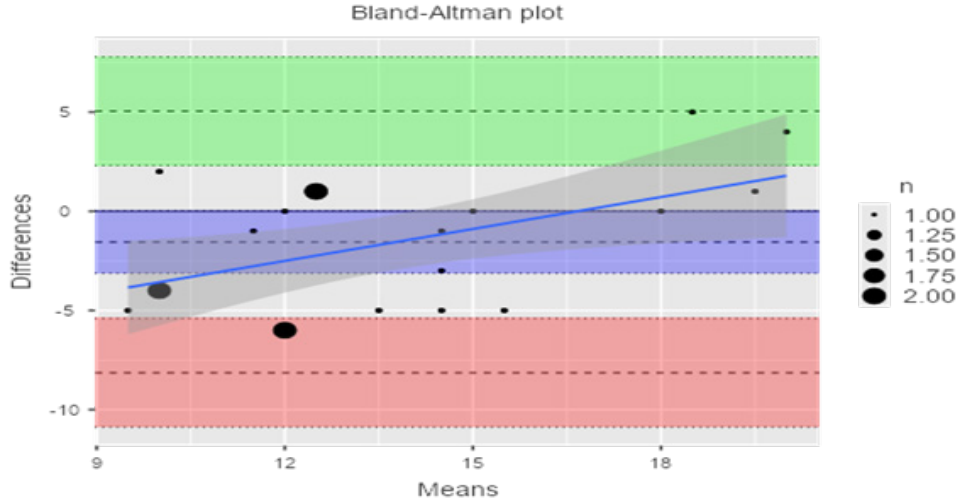
Sağlık alanında yapılan yapay zekâ çalışmalarında CONSORT-AI protokolüne uyum bakımından makalelerde SKK katsayıları bakımında değişimin fazla olduğu belirlenmiştir.

Tablo 2. Türkçeye uyarlanmış CONSORT-AI protokolü için SKK değerleri

	Güven sınırları			
	SKK	Alt Sınır	Üst Sınır	p
1	0,803	0,705	0,901	0,01
2	0,781	0,683	0,879	0,001
3	0,779	0,681	0,877	0,001
4	0,639	0,541	0,737	0,003
5	0,850	0,752	0,948	0,001
6	0,750	0,652	0,848	0,045
7	0,776	0,678	0,874	0,035
8	0,835	0,737	0,933	0,031
9	0,951	0,853	0,956	0,004
10	0,955	0,857	0,960	0,041
11	0,942	0,844	0,947	0,047
12	0,710	0,612	0,808	0,044
13	0,901	0,803	0,999	0,021
14	0,873	0,775	0,971	0,042
15	0,781	0,683	0,879	0,044
16	0,742	0,644	0,840	0,041
17	0,917	0,819	0,922	0,001
18	0,803	0,705	0,901	0,001
19	0,762	0,664	0,860	0,001
20	0,847	0,749	0,945	0,01
21	0,938	0,840	0,943	0,001
22	0,686	0,588	0,784	0,042

Tablo 3. Türkçeye uyarlanmış CONSORT-AI protokolü inceleme sonuçlarının Bland-Altman grafiği için tanımlayıcı istatistikler

CONSORT-AI protokolüne göre makalelerin aldığı puanlar	İlk Uzman		İkinci Uzman		Uzmanlar Arasındaki Fark		
	Ortalama	Değişim Aralığı	Ortalama	Değişim Aralığı	Ortalama	Değişim Aralığı	Güven Aralığı
	13	7;22	14,6	9;19	-1,55	-8;2	-6; 5



Şekil 2. Türkçeye uyarlanmış CONSORT-AI protokolü için uzmanların makalelere verdiği puan ortalamalarına (X Eksen) ve uzmanlar arasındaki farklara (Y Eksen) göre çizdirilen Bland-Altman grafiği

Çalışmalarda net standartlar olmadığı için değerlendiriciler arasında farklılıklar oluşmuştur. Yapılan değerlendirme sonucunda CONSORT-AI protokolünü tümüyle sağlayan bir makaleye rastlanmamıştır.

Tablo 3, makalelere verilen toplam puanların ortalama ve değişim aralıkları, değerlendiriciler arası farkların ortalaması, standart sapması ve güven aralıklarını gösterir. Makalelerin aldığı puanların ortalaması yaklaşık 13,7 olarak hesaplanmıştır. Ayrıca, farklara ilişkin güven aralıkları sıfırı içerdiği için yapılan değerlendirmelerde yanlılığın olmadığı sonucuna varılmıştır.

Şekil 2'deki grafikte değerlendiriciler arası farklar dikkate alınarak Bland-Altman grafiği çizilmiştir. Grafikte CONSORT-AI protokolüne göre yapılan değerlendirmede makalelerin aldığı ortalama puanlar X ekseninde, çalışmalara verilen toplam puanlar bakımından değerlendiriciler arasındaki farklar ise Y ekseninde gösterilmiştir. Türkçeye uyarlanmış CONSORT-AI protokolüne göre makalelerin aldıkları puanların 7-19 arasında değiştiği saptanmıştır.

## Tartışma

EQUATOR Network (Enhancing the Quality and Transparency of Health Research-Sağlık Araştırmalarının Kalitesini

ve Şeffaflığını Artırma Ağı), şeffaf ve doğru raporlamayı ve sağlam raporlama kılavuzlarının daha geniş kullanımını teşvik ederek yayınlanmış sağlık araştırması literatürünün güvenilirliğini ve değerini iyileştirmeyi amaçlayan uluslararası bir girişimdir (EQUATOR Network, 2021). CONSORT 2010, CONSORT-AI, STROBE, SPIRIT, PRISMA, STARD, STROCSS, SAGER, TIDier, CARE, AGREE, SRQR, ARRIVE, SQUIRE, CHEERS gibi raporlama rehberleri/protokolleri, tıp ve sağlık bilimleri alanında yapılan araştırmaların niteliğini artırmanın yanı sıra, söz konusu alanlarda üretilen bilginin standart biçimde sunumunu da teşvik etmektedir. On altı dilde çevirisi yapılan bu rehberlerden CONSORT 2010, STROBE, STARD, STROCSS, SAGER ve TIDier, konuyla ilgili araştırmacıların katılımıyla ve disiplinler arası bir yaklaşımla Türkçe 'ye kazandırılmıştır. Anılan listelerin ortak amacı, yetersiz raporlama sorunlarının sistematik biçimde ve küresel ölçekte çözülmesine yönelik girişimlerde bulunmaktır (EQUATOR Network, 2021). EQUATOR Network sağlık araştırmalarının raporlamasıyla ilgili güncel bilgi ve materyalleri sağlayan kapsamlı bir çevrimiçi kaynak koleksiyonuna sahiptir. Klinisyenler ve araştırmacıların doğru raporlamanın önemine ilişkin farkındalıklarının artırılmasında etkili bir rol üstlenmektedir. (EQUATOR Network, 2021).

Günümüzün önde gelen tıp dergilerinin sağlıkta yapay zekâ uygulamalarının eksiksiz ve anlaşılır şekilde raporlaştırıl-

masını sağlamak için yukarıda sözü edilen protokollerden CONSORT-AI protokolünü yaygın olarak benimsedikleri söylenebilir (İbrahim ve diğerleri, 2021, s. 4). CONSORT-AI bilginin şeffaf bir anlayışla değerlendirilmesini, anlaşılmasını ve yorumlanmasını sağlayarak, tıp ve sağlık alanındaki araştırmacıların, hakemlerin, editörlerin, bilgi profesyonellerinin ve ilgili diğer kullanıcıların bilgiye daha bütüncül ve güvenilir yaklaşmasına olanak yaratmaktadır. Bu durum, ilgili konularda gelişen literatürün güvenilirliğine ilişkin soru işaretlerinin de sayıca azalmasına katkı sağlayabilir. Tıp ve sağlık alanında güvenli bilginin sağlık araştırmacılarına olduğu kadar, bilgiye dayalı hizmetlerin alıcısı konumunda bulunan tüm insanlara yarar sağlayacağı düşünüldüğünde, CONSORT-AI gibi protokollerin gerekliliği daha fazla önem kazanmaktadır. Öte yandan, yapay zekânın hızla gelişen yapısı CONSORT-AI protokolünün güncellenmesine ilişkin gereksinimleri de beraberinde getireceğinden, söz konusu protokolün yeniliklere uyarlanabilir nitelikte gözden geçirilmesi gerekecektir.

CONSORT-AI klinik araştırmalarda yapay zekâ uygulamalarıyla ilgili araştırma çıktılarında açıklanması gereken içerik öğelerini belirleyerek, yazarların yayınlarını daha eksiksiz bir şekilde rapor etmelerine; değerlendirmeye konumunda bulunan hakemlerin ve editörlerin söz konusu yayınları daha şeffaf bir yaklaşımla incelemelerine katkı sağlamaktadır. Ayrıca, yapay zekânın klinik ortama nasıl entegre edildiğini, araştırma verilerinin nasıl elde edildiğini, bulgular doğrultusunda ulaşılan vargıların klinik karar verme sürecine nasıl katkıda bulunduğunu açıkça ortaya koyduğundan, tıp ve sağlık alanındaki bilgi profesyonellerinin işini kolaylaştırma potansiyeli de taşımaktadır.

Bu araştırmada TR Dizi'nde 2012-2020 yıllarında yayınlanan ve yayın dili Türkçe olan 22 makale çerçevesinde CONSORT-AI protokolüne dayalı uyum analizi gerçekleştirilmiştir. Ulaşılan bulgular, CONSORT-AI ölçütlerine bağlı değerlendirmelerde yanlılığın olmadığını, değerlendiriciler arası uyumun anlamlı olduğunu bir başka ifadeyle CONSORT-AI protokolünde yer alan ölçütlerin değerlendiriciler tarafından benzer şekilde algılandığını göstermiştir. Bu durum, CONSORT-AI protokolünün Türkçe uyarlamasının yapay zekâ uygulamaları içeren, Türkçe yazılmış klinik araştırmalarda kullanılabilirliği öne sürülebilir. Öte yandan incelenen yayınlar kapsamında, CONSORT-AI protokolünü tümüyle sağlayan bir makaleye rastlanmamıştır. Bu durum Türkçe uyarlaması yapılan CONSORT-AI'nın konuya ilişkin ulusal literatürün niteliğinin gelişmesine katkı sağlayacağını düşündürmektedir. Yapay zekâ uygulamalarını klinik araştırmalara uyarlama girişimi arttıkça bilimsel araştırma çıktıları etkin, anlaşılır ve daha nitelikli düzeyde raporlandırmaya dönük gereksinim de artacaktır. Bundan hareketle, Türkçe uyarlaması yapılan CONSORT-AI protokolünün Türkiye'deki araştırmacılara, araştırmaları değerlendirenlere, araştırmalara dayalı verileri ve araştırma çıktıları yönetme konumunda bulunanlara yol göstereceği düşünülmektedir. Kanıt değeri taşıyan bilgilerin

şeffaf bir anlayışla düzenlenmesi, nitelikli değerlendirilmesi ve erişime sunulması gibi hizmetlerin belli ilkeler kapsamında gerçekleşmesi, bilginin güvenilirliği konusundaki önyargıları azaltabilir. Bu durum CONSORT-AI protokolünün tıp ve sağlık alanında yaygın kullanımını teşvik eder niteliktedir.

## Öneriler

Protokolleri tıp ve sağlık bilimleri alanında gerçekleştirilen araştırma çıktılarının kalitesini, anlaşılabilirliğini ve şeffaflığını artırıcı niteliktedir. Son dönemde tıp literatüründe hızla yaygınlaşan yapay zekâ odaklı klinik müdahaleler ve bunlara bağlı rapor, dolayısıyla bilgi üretimi, CONSORT-AI gibi protokolleri çerçevesinde gerçekleştirildiği sürece eksiksiz ve standart bir yapıdan söz etmek olanaklı hale gelecektir. Bu kapsamda tıp ve sağlık bilimleri alanındaki araştırmacıların, yayınlarının kalitesini artırmak ve tıpta yapay zekâ uygulamalarına dayalı klinik müdahalelerin şeffaf bir anlayışla raporlandırılmasını sağlamak adına CONSORT-AI destekçisi olmaları beklenmektedir. Öte yandan sürecin bir başka paydaşı olan sağlık alanındaki bilgi profesyonelleri de tıp ve sağlık bilimleri alanındaki araştırma çıktılarının raporlandırılmasında ve dolayısıyla üretilen bilgilerin sunumunda standardizasyonu sağlamak için CONSORT-AI gibi protokollerin kullanımını teşvik etmelidir. Bu çerçevede CONSORT-AI kılavuzuyla ilgili ayrıntıların klinisyenler ve araştırmacılarla paylaşımı gerçekleştirilmelidir.

Bu alanda yer alan değerlendiricilerin ve araştırmacıların süreçlerini şeffaf şekilde yürütebilmeleri için bu gibi kılavuzların kullanımının yaygınlaştırılması gerekmektedir. Elektronik ve yüzyüze eğitimlerin gerçekleştirilmesi; web tabanlı programlar aracılığı ile erişimin sağlanması; klasik eğitim programlarında yer alması ile ilgili tarafların kullanımını arttıracaktır.

## Çıkar Çatışması

Yazarlar herhangi bir çıkar çatışması bildirmemiştir.

## 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 karar olumsuz etkileyecek maddi ve/veya manevi herhangi bir destek alınmamıştır.

## Kaynaklar

Ahuja AS, 2019. The impact of artificial intelligence in medicine on the future role of the physician. Peer J, 7, 1-19.



- Amisha MP, Pathania M, Rathaur VK, 2019. Overview of artificial intelligence in medicine. *Journal of Family Medicine and Primary Care*, 8(7), 2328–2331.
- Ateş C, Öztuna D, Genç Y, 2009. Sağlık araştırmalarında sınıf içi korelasyon katsayısının kullanımı. *Türkiye Klinikleri J Biostat*, 1(2), 59-64.
- Balak T, Günay F, Kaymaz Ö, Çevrimli MB et al., 2018. Reporting observational studies in veterinary research: Turkish adaptation of the STROBE-VET criterias. *Eurasian J Vet Sci*, 34(4), 257-264.
- CONSORT, 2021. Welcome to the CONSORT website. <http://www.consort-statement.org/>, Erişim Tarihi: 25.08.2021
- Densen P, 2011. Challenges and opportunities facing medical education. *Trans Americ Clin Climat Assoc*, 122, 48.
- EQUATOR Network, 2021. About us website. <https://www.equator-network.org/about-us/>, Erişim Tarihi:30.08.2021
- Genç Y, Albayrak R, Ateş C, Tekindal MA et al., 2012. Tanı doğruluğu çalışmalarının kalitelerinin değerlendirilmesi: STARD kriterlerinin Türkçe uyarlaması. *Ankara Üniversitesi Tıp Fakültesi Mecmuası*, 65(3), 137–46.
- İbrahim H, Liu X, Rivera SC, Moher D, Chan et al., 2021. Reporting guidelines for clinical trials of artificial intelligence interventions: The SPIRIT-AI and CONSORT-AI guidelines. *Trials*, 22(11), 1-5.
- Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT et al., 2004. CONSORT Group: Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Annals of internal medicine*, 141(10), 781-788
- Hopewell S, Clarke M, Moher D, Wager E et al., 2008. CONSORT Group: CONSORT for reporting randomised trials in journal and conference abstracts. *The Lancet*, 371(9609), 281-283.
- Lee J, Koh D, Ong CN, 1989. Statistical evaluation of agreement between two methods for measuring a quantitative variable. *Comp bio medi*, 19(1), 61-70.
- Liu X, Cruz RS, Moher D, Calvert MJ et al., 2020. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *Nat Med*, 26, 1364–1374.
- Masic I, Miokovic M, Muhamedagic B, 2008. Evidence based Medicine – new approaches and challenges. *Acta Informatica Medica*, 16(4), 219–225. DOI: 10.5455/aim.2008.16.219-225
- McGraw KO, Wong SP, 1996. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*, 1(1), 30-46.
- Medical Library Association, 2021. Health information profession website. <https://www.mlanet.org/page/profession>, Erişim Tarihi: 18.09.2021
- Sackett DL, 2007. Commentary: Measuring the success of blinding in RCTs: don't, must, can't or needn't? *International journal of epidemiology*, 36(3), 664-665.
- Schmidt ME, Steindorf K, 2006. Statistical methods for the validation of questionnaires. *Met of Inform in Medicine*, 45(4), 409-413.
- Schulz KF, Altman DG, Moher D, 2010. CONSORT 2010 sta-

tement: Updated guidelines for reporting parallel group randomized trials. *Trials*, 11(1), 1-8.

- Serhatlıoğlu S, Hardalac F, 2009. Yapay zekâ teknikleri ve radyolojiye uygulanması. *Fırat Tıp Dergisi*, 14(1), 1-6.
- Sinsky C, Colligan L, Li L, Prigmet M et al., 2016. Allocation of physician time in ambulatory practice: A time and motion study in 4 specialities. *Annals of Internal Med*, 165(11), 753-60.
- Tekindal MA, Ateş C, Yavuz Y, Albayrak R et al., 2015. Tanı doğruluğu çalışmalarının stard kriterlerine göre türk tıp dizini özelinde değerlendirilmesi. *Ankara Üniversitesi Tıp Fakültesi Mecmuası*, 68(2), 49-53.

### Yazar Katkıları

- Fikir/Kavram: Mustafa Agah Tekindal, Melike Tekindal  
Tasarım: Mustafa Agah Tekindal  
Denetleme/Danışmanlık: Melike Tekindal  
Veri Toplama ve/veya İşleme: Mert Demirsöz, Zeynep Özel  
Analiz ve/veya Yorum: Mert Demirsöz  
Kaynak Taraması: Mustafa Agah Tekindal  
Makalenin Yazımı: Mert Demirsöz, Zeynep Özel  
Eleştirel İnceleme: Mert Demirsöz, Zeynep Özel

### Etik Onay

Bu makaledeki sunulan verilerin, bilgilerin ve belgelerin akademik ve etik kurallar çerçevesinde elde edildiği, tüm bilgi, belge, değerlendirme ve sonuçlarının bilimsel etik ve ahlak kurallarına uygun olarak sunulduğuna dair yazarlardan etik beyan alınmıştır.

**CITE THIS ARTICLE:** Tekindal MA, Tekindal M, Demirsöz M, Özel Z, 2022. Klinik çalışmalarda yapay zekâ uygulamalarının raporlanmasında CONSORT-AI protokolünün kullanımı. *Eurasian J Vet Sci*, 38, 2, 66-74

