

STARD 2015 Kriterlerinin Türkçe Uyarlaması; Tanı Doğruluğu Çalışmalarının Raporlanması İçin Bir Kılavuz

Turkish Adaptation of STARD 2015 Criteria; A Guide to Reporting of Diagnostic Accuracy Studies

- Emre DEMİR^a,
- Yasemin YAVUZ^b,
- Can ATEŞ^c,
- Mustafa Agah TEKİNDAL^d,
- Ümran MUSLU^e,
- Engin YILDIRIM^f

^a Biyoistatistik AD,

^b Plastik, Rekonstrüktif ve Estetik Cerrahi AD,
^c Kadın Hastalıkları ve Doğum AD,
Hitit Üniversitesi Tıp Fakültesi,
Çorum, TÜRKİYE

^d Biyoistatistik AD,

Ankara Üniversitesi Tıp Fakültesi,
Ankara, TÜRKİYE

^e Biyoistatistik AD,

Yüzüncü Yıl Üniversitesi Tıp Fakültesi,
Van, TÜRKİYE

^f Biyoistatistik AD,

Selçuk Üniversitesi Tıp Fakültesi,
Konya, TÜRKİYE

ÖZET Amaç: Bu çalışmada, tanı doğruluğu çalışmalarının raporlama kalitesini iyileştirmek ve bir standart oluşturmak amacıyla 2003 yılında yayınlanan ve uluslararası birçok yayıncı tarafından makale yüklerken doldurulması zorunlu hale getirilen 25 maddelik STARD (the standards for the reporting of diagnostic accuracy studies) kriterlerinin 2015 yılında güncellenmiş 30 maddelik yeni versiyonunun Türkçeye uyarlanması amaçlanmıştır. Araştırmanın, Türkçe yayınlanan tanı doğruluğu çalışmaları için standart oluşturulması yoluyla yayın kalitelerinin artırılmasına katkı sağlayacağı düşünülmektedir. **Gereç ve Yöntemler:** STARD 2015 kriterleri, üç yazar tarafından birbirlerinden bağımsız olarak Türkçeye çevrilmiş ve YY'nin liderliğinde her bir madde için görüş birliği sağlanarak son hali verilmiştir. Türkçeye uyarlanmış kontrol listesinin tekrar edilebilirlik düzeyini değerlendirmek amacı ile iki akademisyen hekim 15 tanı doğruluğu çalışmasının kalitelerini değerlendirmiştir. Değerlendiriciler arasındaki uyum için GWET AC1 istatistiği, Sınıf İçi Korelasyon Katsayısı (SKK) ve Bland-Altman yöntemleri kullanılmıştır. **Bulgular:** GWET AC1 istatistiğine göre 21 kriterin değerlendirme sonuçları arasında tam uyuma olduğu, 8 kriterin 0,81-1,00 aralığında, 5 kriterin 0,61-0,80 aralığında olduğu gözlenmiştir. Hekimlerin makalelere vermiş oldukları toplam puanlar arasında SKK değeri 0,871 (%95 GA: 0,616-0,957) olarak bulunmuştur. **Sonuç:** Birçok uluslararası yayıncı tarafından makale yüklerken doldurulması zorunlu hale gelen STARD 2015 kılavuzunun ülkemizdeki dergilerde tanı doğruluğu ile ilgili çalışmalarda kalitenin artırılması için zorunlu bir kılavuz haline gelmesi yakın gelecekte öngörülmektedir.

Anahtar Kelimeler: STARD 2015 kriterleri; tanı doğruluğu çalışmaları; uyarlama

Received: 25.01.2019

Received in revised form: 29.03.2019

Accepted: 01.04.2019

Available Online: 03.04.2019

Correspondence:

Emre DEMİR
Hitit Üniversitesi Tıp Fakültesi,
Biyoistatistik AD, Çorum,
TÜRKİYE/TURKEY
emredemir82@gmail.com

ABSTRACT Objective: This study aims to perform Turkish adaptation of the 30-item new version of the 25-item STARD criteria updated in 2015, which was originally published in 2003 and had to be filled in many international journals while submitting an article with a view to improving the quality of reporting in diagnosis accuracy studies and forming a standard. The study is believed to contribute to the improvements in publication quality of the Turkish publications by forming a standard for diagnosis accuracy. **Material and Methods:** STARD 2015 criteria were translated to Turkish by three authors separately, and a consensus was reached by giving the final version of each item with YY's guidance. Two doctors who are academicians checked the quality of the 15 diagnosis accuracy draft with a view to evaluating its repeatability level. Agreement among the raters was checked using GWET AC1, Intraclass Correlation Coefficient (ICC) and Bland-Altman methods. **Results:** GWET AC1 statistics showed that there was a full match between the evaluation results of the 21 criteria. Eight criteria were between 0.81 and 1.00 and 5 criteria were between 0.61 and 0.80. ICC value was found to be 0.871 in the total scores given to the articles by the doctors. **Conclusion:** It is predicted that in the near future it will be obligatory to fill in the STARD 2015 guide, which became obligatory in many international journals while submitting an article, with a view to increasing the quality of the studies on diagnosis accuracy in the journals in our country.

Keywords: STARD 2015 criteria; diagnostic accuracy studies; adaptation

Teknolojinin gelişmesine paralel olarak her geçen gün hastalıkların taranmasında, tanı koymada, hastalık evrelerinin belirlenmesinde ve izlem amacıyla eskisine göre daha iyi olduğu iddia edilen yeni yöntemler ortaya çıkmaktadır. Bu durum yeni geliştirilen indeks testin (tanı testinin) bir hastalığın varlığını ya da yokluğunu doğrulamada, hastalığın seyri hakkında bilgi vermede veya bazı durumlarda tedaviye verilen yanıtı belirlemek amacıyla bir tanı testi olarak kullanılıp kullanılmayacağını belirlenmesini de beraberinde getirmektedir. Yeni geliştirilen indeks testlerin ayırıcılık gücünü belirlemek veya iki ya da ikiden fazla testin ayırıcılık gücünü karşılaştırmak amacıyla yapılan çalışmalar “tanı doğruluğu” çalışmaları olarak adlandırılır.^{1,2} Radyolojik görüntüleme yöntemleri, patolojik ve histolojik değerlendirmeler, biyokimyasal testler, bakteri kültür testleri ve fizik muayene bulguları tanı testlerine örnek olarak verilebilir.

Tanı testlerinin ayırıcılık gücü, tanı testi ve kesin referans test (altın standart test) sonuçlarının karşılaştırılması yoluyla elde edilen “tanı doğruluk ölçüleri” yardımıyla belirlenir. Tanı testinin ölçüm biçimine bağlı olarak en sık kullanılan tanı doğruluğu ölçüleri Duyarlılık, Seçicilik, Pozitif ve Negatif Tahmini Değerler, Olabilirlik Oranları ve ROC eğrisi altında kalan alan (AUC)’dır. Ayrıca çalışma tasarımına bağlı olarak, önerilen birden fazla tanı testinin performansını karşılaştırmak amacıyla da her geçen gün yeni istatistiksel yöntemler önerilmektedir. Fakat tanı testi çalışmaları incelendiğinde, testlerin ayırıcılık gücünü inceleyen çalışmaların tasarlanması, yürütülmesi ve raporlanmasına yeterince önem verilmediği görülmektedir.^{1,3,4} Tanı testi sonuçlarında karşılaşılabilecek yanlışlık kaynakları olarak; hasta ve kontrol grubunun seçimine ilişkin yanlışlıklar (spektrum, örneklem, kitle, yönlendirme, katılım, görüntü temelli seçim ve izlemden çıkarma yanlışlıkları) ve bilgi edinilmesi ile ilgili yanlışlıklar (kesin olmayan altın standart test, referans veya altın standart testin uygulanmasına yönelik, hastalık ilerleme, iyileşme, analiz ve basım aşaması, evre kayması ve tarama testlerinde görülen özel yanlışlıklar) olmak üzere iki ana başlık altına incelenebilir. Bahsi geçen bu yanlışlık kaynaklarıyla birlikte prevalans ve testin tekrarlanabilirlik düzeyi gibi özellikleri de dikkatlice ele almak önemlidir.⁵

Bunun yanında çalışmaların yöntemi hakkında en önemli bilgilerin eksik raporlanması, sonuçların eksik ve/veya seçilerek verilmesi çalışmaların eleştirel değerlendirmelerini zorlaştırmakta ve araştırma sonuçlarına olan güveni azaltmaktadır. Ayrıca meta analizi çalışmalarının gün geçtikçe artması da, çalışma sonuçlarının raporlanmasında standartların oluşturulması gerekliliğini ortaya koymaktadır. Bu gereklilikten yola çıkarak araştırma tipine bağlı olarak yayınların değerlendirmesi amacıyla kontrol listeleri/kılavuzlar geliştirilmiştir hatta tüm kılavuzların yer aldığı web sayfaları oluşturulmuştur (<http://www.equator-network.org/>). Bunlardan en sık kullanılanlar CONSORT (Paralel Gruplu Randomize Çalışmaların Bildirimi Kılavuzu), STROBE (Gözlemsel Çalışmaların Bildirimi Kılavuzu), PRISMA (Sistemik Derleme ve Meta Analizler için Kılavuz), CARE (Vaka Bildirisi Yazım Kılavuzu), GRRAS (Güvenilirlik ve Uygunluk Çalışmaları Bildirimi Kılavuzu), ARRIVE (Hayvan Çalışmalarının Bildirilmesi Kılavuzu), STARD (Testlerin Tanısal Doğruluklarını İnceleyen Çalışmalar için Yazım Kılavuzu) ’dır.^{3,6-11} Son yıllarda Elsevier, Springer gibi birçok uluslararası yayıncı dergilerinde tanı doğruluğu konusundaki makalelerin incelemeye alınması için kontrol listesinin doldurulmasını zorunlu hale getirmiştir. Bunun yanında uluslararası dergilerde görev yapan hakemler de çalışma kalitesini değerlendirmek için bu kılavuzlardan yararlanmaktadır.

Son yıllarda bu kılavuzların farklı dillere uyarlanmasının yapıldığı birçok makale yayınlanmıştır. CONSORT; Sunay ve arkadaşları tarafından 2013 yılında, PRISMA; Gülpınar ve Güçlü tarafından 2013 yılında, STROBE; Karaçam tarafından 2014 yılında, GRRAS; Erdoğan ve Temel tarafından 2016 yılında Türkçeye uyarlanmıştır.¹²⁻¹⁵

Tanı doğruluğu çalışmalarını değerlendirmek için STARD 2003 kriterleri Genç ve arkadaşları tarafından 2012 yılında Türkçeye uyarlanmıştır.¹⁶ Fakat 2015 yılında STARD kriterleri güncellenmiş ve 30 mad-

deden oluşan kriterler “STARD 2015” adı altında 3 farklı dergide (British Medical Journal, Radiology, Clinical Chemistry) yayınlanmıştır.¹⁷⁻¹⁹ STARD 2015 kriterlerinde toplam 30 madde bulunmakta ancak 4 madde a ve b alt maddelerinden oluştuğu için kılavuz kullanılırken toplam 34 madde değerlendirilmektedir. Araştırmacıların sisteme yükleyecekleri makalelerinin içerisinde her bir kriterin yeterliliğini sağlayan sayfa numarasını kılavuzun son sütununa yazmak suretiyle kılavuzu tamamlamaları gerekmektedir. Güncellenen kılavuza yapılandırılmış özet, testin kullanım amacı ve klinik rolü, çalışma hipotezleri, örneklem büyüklüğü, yapılandırılmış tartışma, kayıt (klinik çalışma kayıt numarası; www.clinicaltrials.gov vb.), protokol ve finansman kaynaklarının açıklandığı yeni maddeler eklenmiştir.¹⁷⁻¹⁹

STARD'ın geliştirilmesindeki temel amaç, araştırmacıların okudukları çalışmalarda ortaya çıkması mümkün olan yanlışlıkları, çalışma bulgularının uygulanabilirliğini ve sonuçların genellenebilirliğini değerlendirmelerine olanak sağlamanın yanında bu konuda makale yazarlara yol göstermektir.^{3,16}

Araştırma kapsamında, 2003 yılında yayınlanan 25 maddelik STARD kriterlerinin güncellenmiş 30 maddelik 2015 versiyonunun Türkçeye uyarlanması amaçlanmaktadır. STARD 2015 kriterlerinin Türkçe yayınlanan tanı doğruluğu çalışmaları için standart oluşturması ve bu yolla yayın kalitelerinin arttırılmasına katkıda bulunması hedeflenmektedir.

GEREÇ VE YÖNTEMLER

DİL EŞDEĞERLİLİĞİ ÇALIŞMASI

STARD kriterlerinin Türkçeye uyarlanması sürecinde öncelikle, orijinal STARD 2015 kriterleri, üç yazar (ED, CA, MAT) tarafından bağımsız olarak Türkçeye çevrilmiştir. Çevrilen bu üç metin tanı doğruluğu çalışmaları konusunda uzman YY'nin liderliğinde görüş birliği sağlanarak son bir metin haline dönüştürülmüştür. Dil eşdeğerliliğinin sağlanması amacıyla, Türkçe uyarlaması yapılan STARD 2015 kriterleri, profesyonel bir tercüman tarafından incelenerek gerekli düzenlemeler yapılmıştır. Son hali verilen kontrol listesi, Türkçe ve İngilizce dillerine eş düzeyde hakim bir başka profesyonel tercüman tarafından Türkçeden İngilizceye çevrilmiş ve orijinal STARD kriterleri ile İngilizceye çevrilmiş STARD kriterleri karşılaştırıldığında aralarında üst düzeyde bir uyum olduğunu saptamıştır.¹⁶

STARD 2015 kriterlerinin Türkçe uyarlaması Tablo 1'de ve uyarlama sürecinde izlenen aşamalar ise Şekil 1'de sunulmuştur. Katılımcıların araştırmaya dâhil edilme kriterlerini gösteren akış diyagramı (STARD 2015 kriterleri 6. madde) Şekil 2'de sunulmuştur.

Türkçeye çevrilen kılavuzun araştırmacılar tarafından doğru bir şekilde anlaşılıp anlaşılmadığını ve tekrar edilebilirlik düzeyini saptamak için 15 tanı doğruluğu çalışması farklı uzmanlığa sahip iki akademisyen hekime (Gözlemci I; ÜM ve Gözlemci II; EY) verilmiş ve Türkçe STARD kontrol listesini kullanarak, yayınların kalitelerini değerlendirmeleri istenmiştir.

YAYIN SEÇİMİ

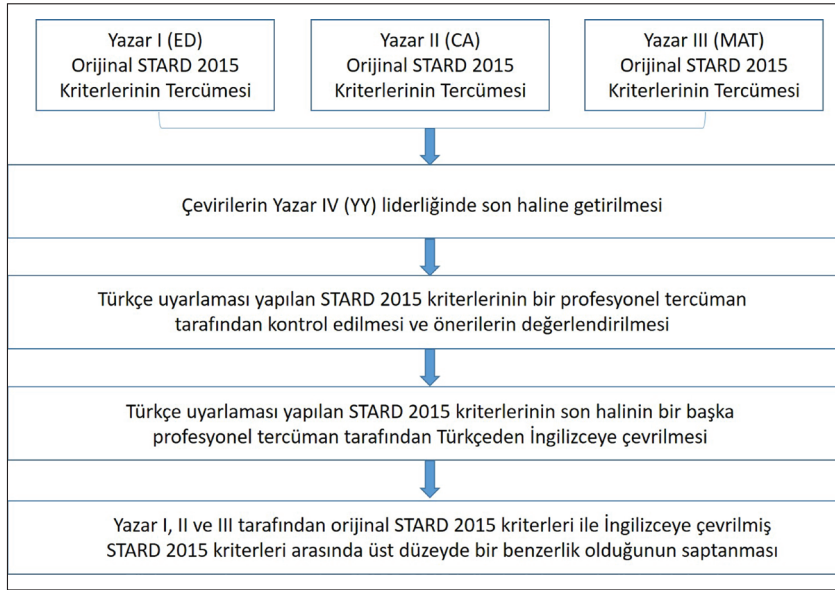
Türkçe STARD 2015 kriterlerinin değerlendirilmesinde kullanılacak yayınların belirlenmesi amacıyla Türk Tıp Dizininde 2015-2016 yıllarında Türkçe dilinde yayınlanmış, insanlar üzerinde yapılmış, başlık, özet ve anahtar kelimeler içerisinde “duyarlılık” ve/veya “seçicilik/özgüllük” ve/veya “tanı doğruluğu/performansı” ve/veya “ROC/İKE/İşlem Karakteristiği Eğrisi” kelimeleri bulunan çalışmalar taranmıştır. Çıkan makaleler içinden 15 orijinal araştırma makalesi ilk dört yazar tarafından belirlenmiştir.

İSTATİSTİK YÖNTEMLER

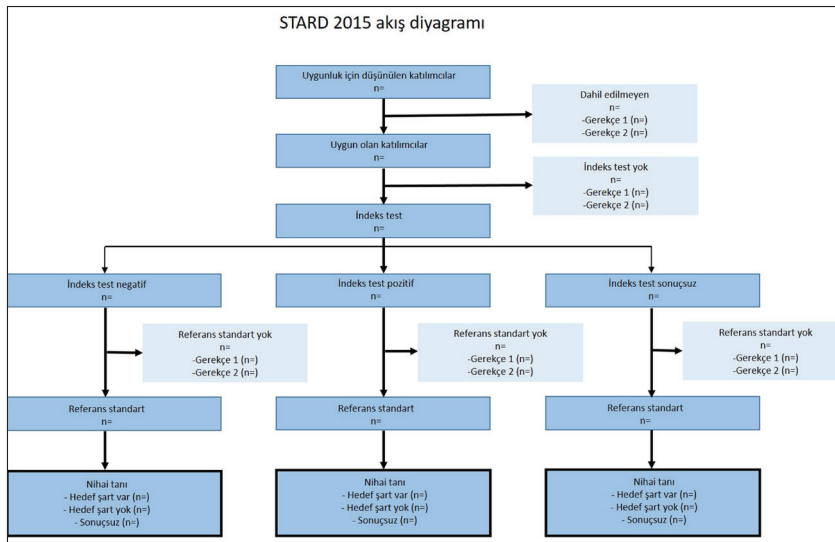
Araştırmamızda istatistik analizler için SPSS (Versiyon 22.0, Lisans: Hitit Üniversitesi) paket programı kullanılmıştır. Bland-Altman grafiği için açık kaynak kodlu R paketi kullanılmıştır. Gözlemciler arasın-

TABLO 1: STARD 2015 kriterlerinin Türkçe uyarlaması.

Bölüm ve Konu	No	Madde	Sayfa Numarası
BAŞLIK veya ÖZET			
	1	En az bir doğruluk ölçüsü (duyarlılık, seçicilik, pozitif-negatif tahmin değerler veya İşlem Karakteristiği Eğrisi (ROC) altında kalan alan (AUC) vb.) kullanarak çalışmanın bir tanı doğruluğu çalışması olarak tanımlanması	
ÖZET			
	2	Çalışma tasarımı, yöntemler, bulgular ve sonuçların yapılandırılmış özeti (bk. STARD Özetler)	
GİRİŞ			
	3	Tanı testinin klinik rolünü ve tasarlanma amacını içeren bilimsel ve klinik arka plan	
	4	Çalışmanın amacı ve hipotezler	
YÖNTEMLER			
<i>Çalışma tasarımı</i>	5	Veri toplama sürecinin indeks test (tanı testi) ve referans standart test (altın standart test) uygulanmadan önce mi (prospektif) yoksa sonra mı (retrospektif) planlanmış olduğu	
<i>Katılımcılar</i>	6	Dâhil etme kriterleri	
	7	Potansiyel olarak uygun bulunan katılımcıların nelere dayanarak çalışmaya dahil edildiği (belirtiler, önceki test sonuçları, anamnez bilgisi)	
	8	Potansiyel olarak uygun bulunan katılımcıların nerede ve ne zaman seçildiği (ortam, konum, tarih bilgisi)	
	9	Katılımcıların araştırmaya seçilme yöntemi (ardışık, rastgele veya ulaşma uygunluklarına göre)	
<i>Test yöntemleri</i>	10a	Tekrar edilebilirliği sağlamak için tanı testinin ayrıntılı açıklaması	
	10b	Tekrar edilebilirliği sağlamak için referans standart testin ayrıntılı açıklaması	
	11	Referans standart testin seçilmesinin gerekçesi (alternatifleri olması durumunda)	
	12a	Araştırmadan önce belirlenmiş, tanı testi sonucu için pozitifliği gösteren eşik değerlerin ve tanı kategorilerinin tanımı	
	12b	Araştırmadan önce belirlenmiş, referans standart test sonucu için pozitifliği gösteren eşik değerlerinin ve tanı kategorilerinin tanımı	
	13a	Tanı testi uygulayan/değerlendirenlerin referans standart test sonuçlarına veya deneklere ait klinik bilgiye sahip olup olmadıkları	
	13b	Referans standart testi uygulayan/değerlendirenlerin tanı test sonuçlarına veya deneklere ait klinik bilgiye sahip olup olmadıkları	
<i>Analizler</i>	14	Tanı doğruluğu ölçülerinin tahmin edilmesi ve karşılaştırmasında kullanılan yöntemler	
	15	Tanı testi ve referans standart test sonucunda tanıda kararsız kalınan olguların nasıl ele alındığı	
	16	Tanı testi ve referans standart test sonuçlarındaki kayıp verilerin nasıl ele alındığı	
	17	Tanı doğruluğunun değişkenliğini belirlemek için analizler (çalışmaya başlamadan önce alt grup analizi planlaması var mı?)	
	18	Hedeflenen örneklem büyüklüğü ve nasıl hesaplandığı	
SONUÇLAR			
<i>Katılımcılar</i>	19	Katılımcılara ilişkin akış diyagramı	
	20	Katılımcıların temel demografik ve klinik özellikleri	
	21a	Hedef koşulu sağlayanlarda hastalık şiddetinin dağılımı	
	21b	Hedef koşulu sağlamayanlarda alternatif tanıların dağılımı	
	22	Tanı testi ile referans standart test uygulamaları arasındaki zaman aralığı ve bu süreçte uygulanan klinik müdahaleler	
<i>Test sonuçları</i>	23	Referans standart testin sonuçlarına göre tanı testinin sonuçlarının (veya dağılımlarının) çapraz tablo ile gösterilmesi	
	24	Tanı doğruluğu kestirimleri ve güvenilirlik değerleri (%95 güven aralıkları gibi)	
	25	Tanı testi veya referans standart testin uygulamasından kaynaklanan yan etkiler	
TARTIŞMA			
	26	Olası yanlışlık kaynakları, istatistiksel belirsizlikler ve genellenebilirlik dâhil olmak üzere çalışmanın kısıtlılıkları	
	27	Tanı testinin amaçlanan kullanımı ve klinik rolünü de içeren uygulamaya yönelik çıkarımlar/öneriler	
DİĞER BİLGİLER			
	28	Kayıt numarası ve kayıt adı	
	29	Çalışma protokolünün tam metnine erişim	
	30	Finansman kaynakları ve diğer destekler; fon sağlayıcıların rolü	



ŞEKİL 1: STARD 2015 kriterlerinin Türkçeye uyarlanma aşamalarını gösteren akış diyagramı.



ŞEKİL 2: STARD akış diyagramı; bir tanı doğruluğu çalışması boyunca katılımcıların bilgisini raporlamak için çalışmanın aşamaları boyunca ilerleme akış diyagramı.

daki doğru değerlendirmeler yüzde olarak sunulmuş olup değerlendiriciler arasındaki uyum için GWET AC1 istatistikleri hesaplanmıştır. Gözlemcilerin makalelere vermiş oldukları STARD toplam puanları arasındaki uyum için Sınıf İçi Korelasyon Katsayısı (SKK) hesaplanmış ve Bland-Altman grafiği kullanılarak görsel olarak değerlendirilmiştir.²⁰

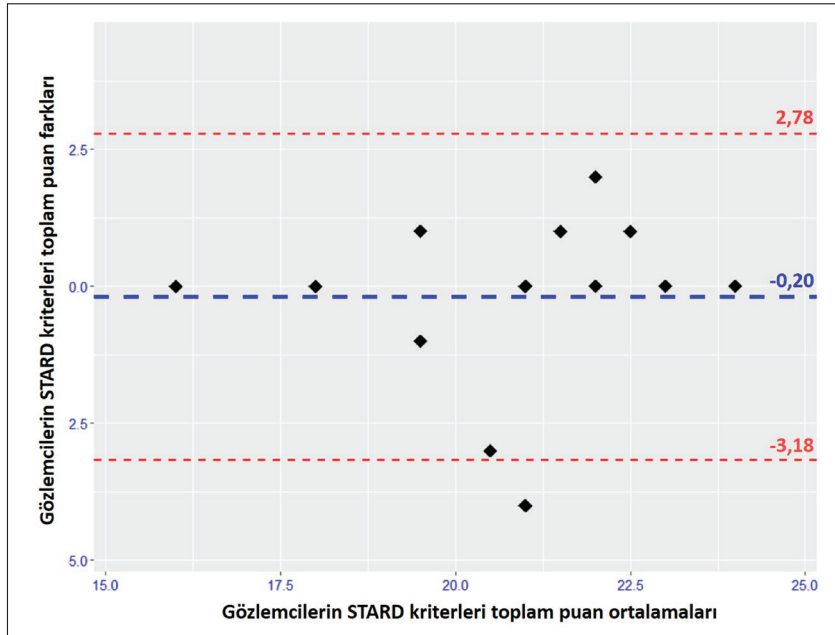
Bu çalışma Helsinki Deklarasyonu 2013 prensiplerine uygun olarak yürütülmüştür.

BULGULAR

Akademisyen gözlemciler tarafından makalelerin STARD 2015 kriterlerine göre değerlendirilmesi sonucunda her bir kriterin raporlanma yüzdesi, gözlemciler arası uyum yüzdeleri, GWET AC1 istatistiği Tablo 2'de sunulmuştur. Tablo 2 incelendiğinde kriterlerin rapor edilme yüzdeleri açısından gözlemciler arasın-

TABLO 2: STARD 2015 kriterlerinin rapor edilme yüzdeleri, her bir kriter için gözlemciler arası uyum yüzdeleri ve GWET AC1 istatistikleri (n=15).

Maddeler	Kriterlerin Rapor Edilme Sayıları		Gözlemciler Arası Uyum Yüzdesi	Uyum istatistikleri
	Gözlemci I (n)	Gözlemci II (n)	%	GWET AC1
1	15	15	100	1
ÖZET				
2	13	13	100	1
GİRİŞ				
3	15	15	100	1
4	15	15	100	1
YÖNTEMLER				
5	12	11	93	0,896
6	14	14	100	1
7	12	14	87	0,827
8	14	14	100	1
9	6	6	87	0,744
10a	15	15	100	1
10b	15	15	100	1
11	10	9	93	0,876
12a	14	14	87	0,848
12b	14	14	87	0,848
13a	14	14	100	1
13b	14	14	100	1
14	14	14	100	1
15	7	7	87	0,735
16	0	0	100	1
17	1	1	93	0,929
18	1	1	100	1
SONUÇLAR				
19	0	0	100	1
20	13	13	100	1
21a	9	9	87	0,744
21b	9	9	87	0,744
22	4	4	87	0,781
23	8	9	93	0,869
24	15	15	100	1
25	0	0	100	1
TARTIŞMA				
26	6	6	93	0,876
27	15	15	100	1
DİĞER BİLGİLER				
28	0	0	100	1
29	0	0	100	1
30	1	1	100	1



ŞEKİL 3: İki farklı gözlemcinin STARD kılavuzuna göre 15 makale içerisinde raporlandığını belirttikleri toplam madde sayılarının uyumuna ilişkin Bland-Altman grafiği.

da toplam 34 kriterden 21 kriterde %100, 5 kriterde %93 ve 8 kriterde ise %87 uyum olduğu gözlenmiştir. GWET AC1 istatistikleri incelendiğinde 21 kriterin değerlendirme sonuçları arasında tam uyuma olduğu, 8 kriterin 0,81-1,00 aralığında, 5 kriterin 0,61-0,80 aralığında olduğu gözlenmiştir.

Üç soruda a ve b seçenekleri olması nedeniyle toplam 34 kriterden Gözlemci I ortalama $20,9 \pm 2,3$ (16-24)'unu raporlamışken, Gözlemci II $21,1 \pm 2,1$ (16-24)'ini raporlamıştır. Makalelerde 34 kriterin en az 16'sı en fazla 24'ü raporlanmış olup, tamamı raporlanan makaleye rastlanmamıştır. Kriter raporlama sayıları toplam puan olarak hesaplanmış ve gözlemcilerin toplam puanları arasındaki SKK değeri 0,871 (0,616-0,957) olarak bulunmuştur. Ayrıca puanlar arasındaki Bland-Altman grafiği Şekil 3'te verilmiştir. Grafik incelendiğinde gözlemcilerin STARD kriterleri toplam puan farklarının sıfırın etrafında rastgele bir dağılım gösterdiği ve bir tanesi hariç ölçüm farklarının %95 güven aralığı olan $-3,18 - 2,78$ ($X - 1,96\sigma - X + 1,96\sigma$) içerisinde bulunduğu görülmektedir.

TARTIŞMA

Genç ve arkadaşları tarafından uyarlanan Türkçe STARD 2003 kriterlerinin uygunluğu iki Patolog hekimin bağımsız olarak 20 makaleyi değerlendirmesi ile test edilmiştir. Bu çalışmada gözlemciler arası uyum yüzdesi %80 ile %100 arasında bulunmuştur. Cohen'in Kappa istatistiğine göre 8 maddede 0,81-1,00, 6 maddede 0,61-0,80, 4 maddede 0,41-0,60, 2 maddede 0,21-0,40 aralığında, 3 madde de ise 0,20 den daha düşük düzeyde uyum raporlanmıştır. Toplam raporlanan STARD kriteri sayısı ortalaması makale başına Gözlemci I için $15,9 \pm 3,2$ ve Gözlemci II için $15,2 \pm 3,6$ olarak neredeyse benzer bulunmuştur. Makalelerde 25 kriterin en az 10'u en fazla 21'i raporlanmış olup STARD kriterlerinin tümünü raporlayan herhangi bir makaleye rastlanmamıştır. Bland-Altman grafiğine göre iki gözlemci arasındaki uyumun kabul edilebilir düzeyde olduğu raporlanmıştır. SKK değerinin 0,945 (%95 GA; 0,852-0,979) olarak hesaplanması sonucunda iki farklı gözlemci değerlendirmesi arasında oldukça yüksek bir uyum bulunduğu ifade edilmiştir.¹⁶

Bu araştırma kapsamında uyarlanan kılavuz kullanılarak iki akademisyen hekim tarafından değerlendirilen makalelere ilişkin sonuçlar arasındaki uyum GWET AC1 katsayılarının değerlendirilmesi sonucunda

yüksek derecede uyumlu olarak bulunmuştur. SKK incelendiğinde sonuçların literatüre benzer şekilde uyum sınırları içerisinde olduğu görülmüştür. Bland-Altman grafiğine göre de her iki gözlemcinin 15 yayın içerisinde raporlandığını belirttiği toplam madde sayısından 14 tanesinin uyum sınırları içerisinde olduğu sadece 1 maddenin uyum sınırları içerisinde yer almadığı görülmüştür.

Ayrıca Plastik ve Rekonstrüktif Cerrahi ve Kadın Hastalıkları ve Doğum alanlarından akademisyen hekimlerin makaleleri değerlendirmesi sonunda STARD 2015 kriterleri hakkındaki görüşleri de sorulmuştur. Akademisyen hekimler genellikle değerlendirdikleri makalelerde yöntem bölümünde bulunması gereken bazı maddelerin tartışma bölümünde verildiğini ifade etmişlerdir. Ayrıca kılavuzda çok fazla sayıda madde bulunduğunu belirtmişlerdir.

Araştırmamızda her iki gözlemci tarafından 15 makale STARD kriterlerine göre incelenmiştir. Kriter sayısının fazla olması sebebiyle 15 makalenin değerlendirilmesi sürecinde akademisyen hekimlerin uzun bir zaman ayırması gerekmiştir. Bu yüzden makale sayısının nispeten az tutulması araştırmamızın kısıtlılığı olarak ifade edilebilir. Ayrıca araştırmamızda değerlendirilen 15 tane doğruluğu çalışmasının belirlenmesinde STARD 2003 ve STARD 2015 kriterlerinin yayım tarihleri dikkate alınmamıştır. Tasarlanacak yeni bir araştırma ile değerlendirilecek tanı doğruluğu çalışmaları STARD 2015 kriterlerinin yayım tarihi öncesi ve sonrası olarak değerlendirilerek araştırmacıların güncellenmiş olan kılavuzu dikkate alıp almadığı araştırılabilir.

Araştırmanın diğer benzer çalışmalardan üstünlüğü ise STARD Türkçe uyarlamasının farklı alanlardaki uzmanlarca kriterler bakımından anlaşılabilirliğinin değerlendirilebilmiş olmasıdır. Literatürdeki uyarlama çalışmalarından sadece Genç ve ark. (2012) tarafından kriter anlaşılabilirliği irdelenmiş o çalışmada da, eş uzmanlık seviyesindeki iki patoloğun uyumuna ilişkin değerlendirme yapılmıştır. Dolayısıyla farklı uzmanlık alanlarından klinisyenlerin uyumuna ilişkin değerlendirmenin yapılması bu çalışmanın farklılığı olarak yorumlanabilir.

SONUÇ

Türkçe uyarlaması yapılmış STARD 2015 kriterlerinin tanı doğruluğu çalışmalarının yapılması ve raporlanması aşamalarında araştırmacılara katkı sağlayacak olup, bunun yanı sıra, makalelerin değerlendirme sürecinde hakemlere, editörlere ve yayınlanmış olan çalışmaları okuyan konu ile ilgili profesyonellere önemli katkı ve kazanımlar sağlayacağı düşünülmektedir. Araştırma raporlarında standartlara uygun ortak ölçütlerin kullanılması, bireysel çalışmaların kalitelerini artırması yanında ilerde yapılacak meta analizi çalışmalarının da kalitelerinin artmasına katkı sağlayacaktır.

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

Yazarlar herhangi bir çıkar çatışması veya finansal destek bildirmemiştir.

Yazar Katkıları

Fikir/Kavram: Emre Demir, Yasemin Yavuz; **Tasarım:** Emre Demir, Yasemin Yavuz; **Denetleme/Danışmanlık:** Emre Demir, Yasemin Yavuz, Can Ateş, Mustafa Agah Tekindal; **Veri Toplama ve/veya İşleme:** Emre Demir, Yasemin Yavuz, Can Ateş, Mustafa Agah Tekindal, Ümran Muslu, Engin Yıldırım; **Analiz ve/veya Yorum:** Emre Demir, Yasemin Yavuz, Can Ateş, Mustafa Agah Tekindal, Ümran Muslu, Engin Yıldırım; **Kaynak Taraması:** Emre Demir; **Makalenin Yazımı:** Emre Demir, Yasemin Yavuz, Can Ateş, Mustafa Agah Tekindal; **Eleştirel İnceleme:** Emre Demir, Yasemin Yavuz, Can Ateş, Mustafa Agah Tekindal, Ümran Muslu, Engin Yıldırım

REFERENCES

1. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Lrwig LM, et al; Standards for Reporting of Diagnostic Accuracy. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Standards for Reporting of Diagnostic Accuracy. Clin Chem.* 2003;49(1):1-6. PMID: 12507953 [[Crossref](#)] [[PubMed](#)]
2. Demir E, Yavuz Y. [Covariate adjustment in receiver operating characteristic curve analysis for PSA in diagnosis of prostate cancer]. *Türkiye Klinikleri J Biostat.* 2016;8(2):162-71. [[Crossref](#)]
3. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Lrwig LM, et al; Standards for Reporting of Diagnostic Accuracy. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem.* 2003;49(1):7-18. PMID: 12507954 [[Crossref](#)] [[PubMed](#)]
4. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool or the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol.* 2006;6:9. PMID: 16519814 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
5. Karaağaoğlu E, Karakaya J, Kılıçkap M. Tanı Testlerinin Değerlendirilmesinde İstatistiksel Yöntemler. 1. Baskı. Ankara: Detay Yayıncılık; 2016. p.272.
6. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340:c332. PMID: 20332509 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
7. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ.* 2007;335(7624):806-8. PMID: 17947786 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
8. Beller EM, Glasziou PP, Altman DG, Hopewell S, Bastian H, Chalmers I, et al; PRISMA for Abstracts Group. PRISMA for abstracts: reporting systematic reviews in journal and conference abstracts. *PLoS Med.* 2013;10(4):e1001419. PMID: 23585737 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
9. Gagnier JJ, Riley D, Altman DG, Moher D, Sox H, Kienle G; CARE Group. The CARE guidelines: consensus-based clinical case reporting guideline development. *Dtsch Arztebl Int.* 2013;110(37):603-8. PMID: 24078847 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
10. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol.* 2011;64(1):96-106. PMID: 21130355 [[Crossref](#)] [[PubMed](#)]
11. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol.* 2010;8(6):e1000412. PMID: 20613859 [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
12. Sunay D, Şengezer T, Oral M, Aktürk Z, Schulz KF, Altman DG, et al; CONSORT Group. [The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials]. *Euras J Fam Med.* 2013;2(1):1-10.
13. Gülpınar Ö, Güçlü AG. [How to write a review article?]. *Turkish Journal of Urology.* 2013;39(Özel Sayı 1):44-8. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
14. Karaçam Z. [The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies]. *Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi.* 2014;17:1.
15. Erdoğan S, Orekici Temel G. [Turkish adaptation of the guidelines for reporting reliability and agreement studies (GRRAS) for clinical investigators]. *Cu-kurova Med J.* 2016;41(4):613-21. [[Crossref](#)]
16. Genç Y, Albayrak R, Ateş C, Tekindal MA, Selvi P, Perçinel S, et al. [Evaluation of quality of diagnostic accuracy studies: Turkish adaptation of STARD criteria]. *Ankara Üniversitesi Tıp Fakültesi Mecmuası.* 2012;65(3):137-46.
17. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *BMJ.* 2015;351:h5527. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
18. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al; STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology.* 2015;277(3):826-32. PMID: 26509226 [[Crossref](#)] [[PubMed](#)]
19. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. An updated list of essential items for reporting diagnostic accuracy studies. *Clinical Chemistry.* 2015;61(12):1446-52. [[Crossref](#)] [[PubMed](#)]
20. Ateş C, Öztuna D, Genç Y. [The use of intraclass correlation coefficient (ICC) in medical research: review]. *Türkiye Klinikleri J Biostat.* 2009;1(2):59-64.