

Tablo 1. STARD kriterlerinin Türkçe uyarlaması

Bölüm ve Konu	Madde		Sayfa Numarası
BAŞLIK/ÖZET/ ANAHTAR KELİMELE	1	Makaleyi bir tanı doğruluğu çalışması olarak tanımlayın (MeSH başlığının duyarlılık ve seçicilik olması önerilir).	
GİRİŞ	2	Tanı doğruluğunun tahmin edilmesi, tanı testlerinin karşılaştırılması ya da tanı doğruluğunun alt gruplarda karşılaştırılması gibi ifadeler kullanarak çalışmanın amacını ve/veya hipotezlerini belirtin.	
METOD			
<i>Denekler</i>	3	Çalışma popülasyonunu tanımlayın: Dahil etme ve çıkarma kriterleri, verilerin toplandığı ortam ve mekanlar.	
	4	Katılımcıların çalışmaya alınma prosedürünü tanımlayın: Katılımcı alımı, var olan semptomlara mı, önceki testin sonuçlarına mı, yoksa referans standart (altın standart test) ve/veya indeks test (tanı testi) sonuçlarına mı dayanmaktadır?	
	5	Katılımcı örneklemini tanımlayın: Madde 3 ve 4 deki seçilme kriterlerini sağlayan katılımcılar, çalışmaya ardışık olarak mı alındı? Eğer değilse katılımcıların nasıl seçildiğini belirtin.	
	6	Verilerin toplama biçimini tanımlayın: Veri toplama biçimi indeks test (tanı testi) ve referans standart (altın standart test) uygulanmadan önce mi (ileriye dönük planlanmış çalışma) uygulandıktan sonra mı (geriye dönük planlanmış çalışma) belirlendi?	
<i>Test Yöntemleri</i>	7	Referans standardı (altın standart testi) ve bilimsel dayanağını tanımlayın.	
	8	Materyal ve metodun teknik özelliklerini, ölçümlerin ne zaman ve nasıl yapıldığını da belirtilerek detaylı olarak açıklayın ve/veya indeks test (tanı testi) ve referans standart (altın standart test) için referanslar gösterin.	
	9	İndeks test (tanı testi) ve referans standart (altın standart test) için kullanılan birimleri, kesim noktalarını ve belirlenen kategorilerin geçerliliğini gerekçeli olarak tanımlayın.	
	10	İndeks testi (tanı testini) ve referans standardı (altın standart testi) uygulayan kişilerin sayısını, eğitimini ve uzmanlık düzeyini belirtin.	
	11	İndeks testi (tanı testini) ve referans standardı (altın standart testi) uygulayan gözlemcilerin, katılımcıların klinik bilgilerine ve diğer test sonuçlarına kör olup olmadıklarını belirtin.	
<i>İstatistiksel Yöntemler</i>	12	Tanı doğruluğu ölçütlerinin hesaplanmasında veya karşılaştırılmasında kullanılan yöntemleri ve belirsizlik tahmini (örneğin, %95 güven aralığı) için kullanılan istatistiksel yöntemleri belirtin.	
	13	Eğer yapılmış ise testin tekrar edilebilirliğini hesaplamak için kullanılan yöntemleri belirtin.	
SONUÇLAR			
<i>Katılımcılar</i>	14	Çalışmanın yapıldığı zamanı, katılımcı alımına başlama ve katılımcı alımını sonlandırma tarihleri ile birlikte belirtin.	
	15	Çalışma popülasyonunun klinik ve demografik özelliklerini (örneğin, yaş, cinsiyet, var olan semptomların spektrumu, eşlik eden hastalıklar, uygulanan tedaviler, tedavi merkezleri) raporlayın.	
	16	İndeks test (tanı testi) ve/veya referans standart (altın standart test) uygulanıp/uygulanmayıp, dahil etme kriterlerini sağlayan katılımcıların sayısını bildirin. Katılımcıların neden her iki testi de almadıklarını tanımlayın (akış diyagramı önerilir).	
<i>Test sonuçları</i>	17	İndeks testler (tanı testleri) ile referans standart (altın standart test) arasında geçen zamanı ve bu süreçte herhangi bir tedavi uygulanıp uygulanmadığını belirtin.	
	18	Hedef koşullara sahip olanlarda hastalık şiddetinin (kriter tanımla) dağılımı ve hedef şartlara uymayan katılımcılara ait diğer teşhisleri raporlayın.	
	19	İndeks testlerin (tanı testlerinin) sonuçlarını (belirlenemeyen ve kayıp sonuçlar da dahil olmak üzere) referans standardın (altın standart testin) sonuçları ile karşılaştırarak tablo halinde verin; sürekli sonuçlar için ise test sonuçlarının referans standart (altın standart test) sonuçlarına göre dağılımını verin.	
	20	İndeks test (tanı testi) veya referans standart (altın standart test) uygulanırken karşılaşılan olumsuz durumları raporlayın.	
<i>Tahminler</i>	21	Tanı doğruluğu tahminlerini ve istatistiksel belirsizlik ölçütlerini (örneğin, %95 güven aralığı) raporlayın.	
	22	İndeks testlerde (tanı testlerinde) şüpheli sonuçların, eksik (kayıp) yanıtların ve sapan değerlerin nasıl ele alındığını raporlayın.	
	23	Eğer yapıldıysa, katılımcıların alt grupları, gözlemciler veya merkezler arasında tanı doğruluğunun değişkenlik tahminlerini raporlayın.	
	24	Eğer yapıldıysa, testlerin tekrarlanabilirlik tahminlerini raporlayın.	
TARTIŞMA	25	Çalışma bulgularının klinik uygulanabilirliğini tartışın.	

STARD

The literal translation of the STARD Criteria is given on page 140 of the publication below. The following steps have been followed for verbatim translation.

Genç Y., Albayrak R., Ateş C., **Tekindal M.A.**, Selvi P., Perçinel S., Ceyhan K., Tanı Doğruluğu Çalışmalarının Kalitelerinin Değerlendirilmesi: STARD Kriterlerinin Türkçe Uyarlaması, Ankara Üniversitesi Tıp Fakültesi Mecmuası 2012, 65 (3), 137-145

Language Equivalence Study

In the process of adapting STARD criteria to Turkish, the original STARD statement was independently translated into Turkish by the first four authors (RA, MAT, CA, PS). During the translation phase, the article published by the STARD group in 2003 and explaining the items with examples was used (9, 10). Four different translation texts were finalized by reaching a consensus under the leadership of YG, who is an expert in planning diagnostic accuracy studies. In order to ensure language equivalence, the STARD criteria, which were adapted to Turkish, were checked by a professional translator and the necessary arrangements were made. The finalized checklist was translated from Turkish into English by another professional translator with a peer command of Turkish and English, and when the original STARD criteria were compared with the STARD criteria translated into English, a high level of similarity was found between them.

http://cms.galenos.com.tr/Uploads/Article_18494/AUTFM-65-136-En.pdf

http://cms.galenos.com.tr/Uploads/Article_18353/AUTFM-68-49-En.pdf

Tanı Doğruluğu Çalışmalarının Kalitelerinin Değerlendirilmesi: STARD Kriterlerinin Türkçe Uyarlaması

Evaluation Of Quality Of Diagnostic Accuracy Studies: Turkish Adaptation Of STARD Criteria

Yasemin Genç¹, Rabia Albayrak², Can Ateş¹, Mustafa Agah Tekindal³, Pınar Selvi⁴, Sibel Perçinel⁵, Koray Ceyhan⁶

¹ Ankara Üniversitesi Tıp Fakültesi Biyoistatistik AD

² Ankara Üniversitesi Ziraat Fakültesi Zootekni Bölümü

Biyometri ve Genetik AD

³ Basket Üniversitesi Tıp Fakültesi Biyoistatistik AD

⁴ Cukurova Üniversitesi Tıp Fakültesi Biyoistatistik AD.

⁵ Ankara Üniversitesi Tıp Fakültesi Tıbbi Patoloji AD.

⁶ Ankara Üniversitesi Tıp Fakültesi Tıbbi Patoloji AD. Sitopatoloji BD.

Amaç: Bu çalışmada, tanı doğruluğu çalışmalarının planlanması ve sonuçlarının raporlanması konusunda bir standardın oluşturulması amacıyla çoğunluğunu epidemiyolog ve biyoistatistikçilerin oluşturduğu bir grup tarafından 2003 yılında yayınlanan STARD (The STAndards for Reporting of Diagnostic Accuracy) kriterlerinin Türkçeye çevrilmesi ve ülkemizde yapılan çalışmalarda kullanımının yaygınlaştırılması amaçlanmıştır.

Gereç ve Yöntem: 2003 yılında yayınlanan orijinal STARD kriterleri, ilk 4 yazar (RA, MAT, CA, PS) tarafından birbirlerinden bağımsız olarak Türkçeye çevrilmiş ve YG'nin liderliğinde görüş birliğine varılarak son haline getirilmiştir. Dil eşdeğerliğinin sağlanması amacıyla, Türkçe uyarlaması yapılan STARD kriterleri profesyonel bir tercüman tarafından kontrol edilmiş ve gerekli düzeltmeler yapılmıştır. Son hali verilen Türkçe STARD kriterleri, geri çevirme yöntemi kullanılarak bir başka profesyonel tercüman tarafından Türkçeden İngilizceye çevrilmiş ve orijinal metin ile üst düzeyde bir benzerlik olduğu saptanmıştır. Daha sonra Türkçeye uyarlanmış kontrol listesinin araştırmacılar tarafından doğru bir şekilde anlaşılıp anlaşılmadığını ve tekrar edilebilirlik düzeyini değerlendirmek amacı ile aynı uzmanlık düzeyine sahip iki patoloğa (SP, KC) 20 tanı doğruluğu çalışması verilmiş ve Türkçe uyarlaması yapılmış olan STARD kriterlerini kullanarak, yayınların kalitelerini değerlendirmeleri istenmiştir. Gözlemciler arasındaki uyum her bir kriter için gözlemciler arası uyum yüzdesi ve Cohen'in Kappa istatistiği kullanılarak incelenmiştir. Uyum ayrıca toplam raporlanan kriter sayısı kullanılarak Bland-Altman yöntemi ve Sınıf İçi Korelasyon Katsayısı (SKK) kullanılarak da değerlendirilmiştir.

Bulgular: İncelenen tanı doğruluğu çalışmalarında kriterler rapor edilme sıklıkları bakımından büyük bir varyasyon göstermektedir. Kriterlerin bir bölümü bütün makalelerde rapor edilirken bir bölümü ise neredeyse hiç rapor edilmemiştir. Makale başına toplam raporlanan kriter sayısı 15.9±3.2 (Gözlemci 1'in sonucu) olup 25 kriterin en az 10'unun, en fazla 21'inin raporlanmış olduğu gözlenmiştir. Kriterler için ayrı ayrı hesaplanan gözlemciler arası uyum yüzdeslerinin bir kriter için %60 değerleri için %75 ve üzerinde olduğu gözlenmiştir. İki kriter için Kappa istatistiği elde edilememiş, 3 kriter için önemsiz düzeyde bulunmuş, 20 kriter için ise 0.32'nin üzerinde bulunmuştur. %95 Uyum Limitleri -2.2-3.5 olarak bulunmuş olup bu sonuç iki gözlemci arasındaki uyumun kabul edilebilir düzeyde olduğunu göstermiş, bunun yanında SKK da 0.945 gibi oldukça yüksek bir değer olarak bulunmuştur.

Sonuç: Türkçe uyarlaması yapılmış STARD kriterlerinin ülkemizde tanı doğruluğu çalışması yapacak olan araştırmacıların yanında dergi hakem ve editörlerin için de bir kılavuz olması umut edilmektedir. Çalışmalarda metodolojik standartlara uygun ortak bir ölçüt kullanılması, bireysel çalışmaların yanında meta analizi çalışmalarının da kalitelerinin artmasını sağlayacaktır.

Anahtar Sözcükler: *STARD kriterleri, Tanı doğruluğu çalışmaları*

Aim: The aim of this study is translate the STARD (The STAndards for Reporting of Diagnostic Accuracy) criteria which were developed in 2003 by a group of epidemiologists and biostatisticians in order to establish a standard for planning and reporting the results of the diagnostic accuracy studies.

Materials and Design: The STARD criteria that were published in 2003 have firstly been translated into Turkish language by the first four authors (RA, MAT, CA, and PS) independently and afterwards four translated versions have been combined with consensus of the whole group headed by YG. For assuring the language equivalency between the original and the translated STARD statements, the translated criteria was also checked by a professional interpreter and necessary corrections were made. The final version of the translated criteria is translated into English by another professional interpreter using reverse-translation method and found to be quite similar with the original. For checking the understandability and repeatability of the translated version of the STARD criteria, 2 pathologists were asked to use the translated STARD criteria and evaluate the quality of 20 diagnostic accuracy studies. The inter-rater agreement for each criterion was investigated by using consistency percentages and Cohen's kappa statistic. Moreover, the agreement is also assessed by using Bland-Altman method and Within Group Correlation by using the total number of reported criteria.

Results: The investigated 20 diagnostic accuracy studies were found to be variable with respect to reported criteria percentages. Some of the criteria were found to be reported in all studies whereas some other criteria were reported in none of the studies. Total number of reported criteria per article is found to be 15.9±3.2 (Rater 1) with a minimum of 10 and a maximum of 21. Agreement percentages per criteria were found to be 60% for one of the criterion, and 75% or above for the others. Kappa statistic could not be estimated for two items, was found to be not significant for 3 items and was above 0.32 for 20 items. The limits for agreement were found as -2.2-3.5, and it shows that the inter-rater agreement has an acceptable level along with SKK is calculated to be 0.945.

Conclusion: It is expected that the translated STARD criteria will be a guideline for the researchers in the diagnostics accuracy studies as well as for the journal reviewers and editors. The use of Turkish translated version of STARD as a standard measure will increase the quality of diagnostic accuracy studies as well as the meta analysis studies.

Key Words: *STARD statements, Diagnostic accuracy studies*

İndeks testler (tanı testleri), hasta ve sağlıklı bireylerin oluşturduğu heterojen bir kitlede bireylerin gerçek durumunu (gerçekten hasta olup olmadıklarını) ortaya çıkarmak amacıyla kullanılır.

Doğruluğu kesin olarak kanıtlanmış referans standartlar (altın standart test) ile bireylere “kesin hasta” ya da “kesin sağlıklı” tanısı koyulabilir. Fakat bu testlerin uygulanmalarının zor,

Geliş Tarihi: 11.04.2012 • Kabul Tarihi: 10.02.2014

İletişim

Dr. Can Ateş

Tel : 0 312 595 81 39

E-posta : can.ates@gmail.com

Ankara Üniversitesi Tıp Fakültesi Biyoistatistik AD, Morfoloji Kampüsü Sıhhiye, Ankara

Tanı Doğruluğu Çalışmalarının Stard Kriterlerine Göre Türk Tıp Dizini Özelinde Değerlendirilmesi

Use of the Stard Statement in Reporting Diagnostic Test Research in Turkish Medical Directory

Mustafa Agah Tekindal¹, Can Ateş², Yasemin Yavuz², Rabia Albayrak³, Pınar Selvi³

¹ Department of Biostatistics and Medical Informatics, Faculty of Medicine, Izmir University

² Department of Biostatistics, Faculty of Medicine, Ankara University

³ Department of Animal Science Biometry and Genetics, Faculty of Agriculture, Ankara University

Amaç: Tanı testi çalışmaları incelendiğinde, testlerin ayrıcalık gücünü değerlendiren çalışmaların planlanması, yürütülmesi ve sonuçlarının raporlanmasına gereken önemin verilmediği görülmektedir. Son yıllarda meta analizi çalışmalarının artması, çalışma sonuçlarının raporlanmasında standartların oluşturulması gerekliliğini ortaya koymuştur. Bu gerekliliklerden yola çıkarak 2003 yılında bu standartları oluşturan QUADAS ve STARD isimleri ile iki kontrol listesi geliştirilmiş ve yayınlanmıştır. Genç ve arkadaşları 2012 yılında STARD kriterlerinin Türkçe uyarlamasını yayınlamışlardır. Çalışmamızda Türkçe STARD Kontrol listesi kullanılarak Türkçe yayınlanan tanı doğruluğu çalışmalarının kaliteleri değerlendirilecek ve her bir kriterin sağlanma oranları yıllara göre belirlenecektir.

Gereç ve Yöntem: Çalışmamızda “duyarlılık, özgüllük, seçicilik, işlem karakteristiği eğrisi, İKE, ROC ve tanı doğruluğu” anahtar kelimeleri kullanılarak 2007-2010 yılları arasında ULAKBİM Tıp Veri Tabanında [Türk Tıp Dizini] yayınlanmış tanı doğruluğu çalışmaları taranmış ve 459 makale belirlenmiştir. Daha sonra bu çalışmalar belirlenen kriterlere göre tekrar değerlendirilerek 75 çalışma seçilmiştir. Bu çalışmalar Türkçe STARD kontrol listesine göre değerlendirilerek her bir kriter için sağlanma yüzdeleri bulunmuştur.

Bulgular: STARD kriterlerini en çok sağlayan makaleler 25 maddenin 21’ini karşılarken kriterlere en az uyum makalelerde bu sayı 6’ya düşmüştür. Bunun yanı sıra, 75 makalenin %9’u STARD beyanına %50’den daha az uyum gösterirken, makalelerin %81’i, %50 - %75 arasında bir uyum sergilemektedir. Çalışmaların büyük bölümünde tanı doğruluğuna ilişkin ölçüler verilirken (%91), sadece %32’sinde bunlara ilişkin Güven Aralıkları verilmiştir. Değerlendirilen makalelerde, çalışmaya alınan deneklere ilişkin kriterler genel olarak yüksek oranda karşılanırken (ortalama %84) bu oranlar yıllar içinde de benzerdir. Değerlendirilen çalışmaların %96’sında sonuçların klinik uygulanabilirliğini tartışılmıştır. Bu kriter en çok sağlanan kriterdir.

Sonuç: 2007-2010 yılları arasında ULAKBİM Tıp Veri Tabanında [Türk Tıp Dizini] yayınlanmış tanı doğruluğu çalışmalarının STARD kriterlerini sağlama oranlarının çok yüksek düzeyde olamasa da yeterli olduğu söylenebilir. STARD kontrol listesi Türkçe versiyonunun kullanıma sunulması Türkçe yayınlanan tanı doğruluğu çalışmalarının standarda kavuşturulması için çok önemli bir araç olarak kullanılabilir. Böylece, Türkçe literatürdeki tanı doğruluğu çalışmalarının niteliğinde daha fazla iyileşme sağlanabilir.

Anahtar Sözcükler: **STARD Kriterleri, Tanı Doğruluğu Çalışmaları, Sistematik Derleme**

Aim: When diagnostic testing studies are examined, it is seen that due importance is not attached to planning and conducting studies that evaluate the discriminating power of tests and reporting the results of such studies. Recent increase in the number of meta-analysis studies has indicated the need to create standards for the reporting of results. Based on this need, two statements making up such standards were developed and published in 2003 under the names of QUADAS and STARD. STARD criteria were adapted to Turkish by Genç et al. in 2012. The aim of our study is to identify the quality of the diagnostic accuracy studies published in Turkish Medical Directory by using the Turkish version of the STARD statement and to determine the rate of satisfaction of each criterion by year.

Materials and Methods: To find appropriate studies, the articles that were published on ULAKBİM Medical Database [Turkish Medical Directory] between 2007 and 2010 and included the words, “sensitivity, specificity, discrimination receiver operating characteristic, ROC, diagnostic accuracy” were searched, and 459 articles were determined. As a result of the electronic search, 459 articles were determined. Then those studies were re-evaluated based on the specified criteria. Finally, 75 appropriate articles were chosen. Those studies were evaluated according to the Turkish STARD statement, and percentage of satisfaction was found for each criterion.

Results: The most compliant articles with the STARD statement satisfied 21 items out of 25 whereas the least compliant articles satisfied only 6 items. Moreover, 9% of 75 articles have less than 50% relevance to the STARD statement while 81% of 75 articles are in 50-75% relevance range. A great majority of the studies (91%) provide measures related to diagnostic accuracy, but only 32% give confidence intervals regarding them. In the evaluated articles, the criteria concerning the subjects included in the studies were mostly satisfied at a high rate (84% in average). These rates of satisfaction are similar over the years. In 96% of the evaluated studies, the clinical applicability of the results was discussed. This is the criterion satisfied most.

Conclusions: To conclude, the rate of satisfaction of the STARD criteria by the diagnostic accuracy studies published on the ULAKBİM Medical Database [Turkish Medical Directory] between 2007 and 2010 is adequate, though not very high. The introduction of the Turkish version of the STARD statement may serve as a very important tool for the standardization of the diagnostic accuracy studies published in Turkish. This may lead to an improvement in the quality of the diagnostic accuracy studies in the Turkish literature.

Key Words: **STARD Statements, Diagnostic Accuracy Studies, Systematic Review**

Geliş tarihi : 16.06.2015 • Kabul tarihi: 12.09.2015

İletişim

Arş. Gör. Mustafa Agah TEKİNDAL

Tel: +90 (232) 246 49 49

GSM:+90 555 399 05 99

E-posta: matekindal@gmail.com

Izmir Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı
35350 İzmir Türkiye

İndeks testler (tanı testleri), hasta ve sağlıklı bireylerin oluşturduğu heterojen bir kitlede bireylerin gerçek durumunu (gerçekten hasta olup olmadıklarını) ortaya çıkarmak amacıyla kullanılır. Doğruluğu kesin olarak ka-

nitlanmış referans standart (altın standart test) testler ile bireylere “kesin hasta” ya da “kesin sağlıklı” tanısı konulabilir. Fakat bu testlerin uygulanmalarının zor, maliyetlerinin yüksek ve bazı hastalıkların oluşmasında